January 1980

Use of Extreme Value Theory in Estimating Flood Peaks from Mixed Populations

Ronand V. Canfield

D. R. Olsen

R. H. Hawkins

T. L. Chen

USE OF EXTREME VALUE THEORY IN ESTIMATING FLOOD PEAKS FROM MIXED POPULATIONS

by

R. V. Canfield
D. R. Olsen
R. H. Hawkins
T. L. Chen

The work upon which this publication is based was supported in part by funds provided by the Office of Water Research and Technology (B-173-UTAH, 14-34-0001-9100) U.S. Department of the Interior, Washington, D.C., as authorized by the Water Research and Development Act of 1978.

HYDRAULICS AND HYDROLOGY SERIES
UWRL/H-80/01

Utah Water Research Laboratory
College of Engineering
Utah State University
Logan, Utah 84322

February 1980

ABSTRACT

The flood magnitude for a given frequency or return period is estimated by fitting a probability distribution to the historical annual flood series. The log-Pearson type III distribution has been selected by the Water Resources Council for general use by the federal government, but practitioners should examine an annual flood series and use alternative distributions where they will produce better estimates. Empirical goodness of fit is one criterion for choosing a distribution, but the reasonableness of the assumptions theoretically associated with the form of the distribution should also be considered.

In theory, extreme-value distributions are particularly applicable to flow series composed of the largest flow from each year of record. The Fisher-Tippett extreme-value function, commonly called the Gumbel distribution, has been widely used for flood frequency analysis, but it was found empirically inferior to the log-Pearson type III distribution by the Water Resources Council. The Gumbel is, however, only one of three alternative extreme-value functions, and these have not been systematically investigated for applicability.

All three are examined herein, and plotting tests are provided for making a selection. The generally most appropriate was found to be not the Gumbel distribution, which assumes neither an upper nor a lower bound to the possible flood flows, but rather a form adding a third parameter as an upper bound to the flood flow. The existence of such an upper bound seems reasonable hydrologically, and a maximum likelihood fit of this distribution to 14 stations around the world with over 50 years of record compares favorably with that with the log-Pearson type III distribution. More efficient parameter estimating techniques are, however, needed.

The plotting tests for many series were found to exhibit a break between two linear portions suggesting that the recorded flows may in fact be drawn from two or more populations. The form of a distribution of a series drawn as a mixture from two populations is shown theoretically to be multiplicative with respect to the two functions (rather than having the more commonly used additive form). A five parameter distribution was applied to 11 long-term sequences shown by the plotting test to originate from nonhomogeneous sources. The fit was generally excellent.

ACKNOWLEDGMENTS

The authors wish to acknowledge the cooperation extended by B. B. Bobee and R. Robitaille in providing the data used in this study.

We also express appreciation to the Office of Water Research and Technology of the U.S. Department of the Interior for partial support of this research, OWRT No. B173-Utah, Agreement No. 14-34-0001-9100.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The central relationship for flood control and floodplain management planning is that between peak flow and return period. The relationship is established by selecting an appropriate distribution to represent the population of peak flows, one from each year of record (the annual flood series), and estimating parameters for that distribution that best fit the recorded data.

The primary criterion used to select an appropriate distribution has been goodness-of-fit as measured empirically. Accordingly, the parameters of several distributions are estimated from the same data set. Some goodness-of-fit criterion is then used to choose the best-fitting distribution (e.g., Bobee and Robitaille 1977). The log Pearson type III distribution was selected for general use on federal water resources studies (U.S. Water Resources Council 1976, Appendix 14) on this basis.

The Monte Carlo experiment described in the next section illustrates that serious estimating errors may arise if the distribution is selected solely on the basis of goodness of fit. The magnitudes of these errors clearly demonstrate that empirical fit alone does not provide an adequate basis for selecting a distribution. Theory provides supplemental information. The annual flood event is the maximum or extreme value of all the events occurring during the year; therefore, extreme value theory would seem to provide a reasonable theoretical base to explore and is examined here. Although extreme value distributions have been used in hydrology, no systematic examination of the theory to determine the most appropriate form is reported in the literature.

The first section of this report presents the problem encountered when empirical fit alone is used to select a "best" distribution. The second section deals with application of extreme value theory to stream flows which have homogeneous sources. The results clearly demonstrate the usefulness of extreme value theory. The third section extends extreme value theory to the case in which the events in the annual series are random variables from two different populations (e.g., thunderstorm and cyclonic events). The fourth section describes how one goes about the mechanics of applying these results in flood frequency analysis.

EMPIRICAL FIT

The problem encountered when empirical fit is the sole criterion used to select a "best" distribution to describe a population increases as one uses the distribution to estimate the frequency of rarer events. It is sometimes suggested that no distribution is perfect; therefore, several may do an adequate job, and certainly the "best" fit will be close. This argument may be valid when the distributions are used to estimate probabilities or return periods for frequently occurring events. However, when estimates are needed for extreme or rare events, serious errors can result from use of a distribution selected on the basis of empirical fit because the probabilities of rare events are computed from the tails of a distribution, whereas empirical fit is dominated by the body of the data set. The following Monte Carlo experiment was performed to provide some idea of the magnitude of the problem.

Twenty random samples, each containing 25 values, were generated from a Weibull population with cumulative distribution function

$$F(x) = \begin{cases} 1 - \exp[-(x/30)^b] & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The gamma distribution is considered close to the Weibull (Hager, Bain, and Antle 1971) and is a likely alternative for fitting such data. Both gamma and Weibull distributions were fit to the data sets. The method of White (1969) was used to estimate Weibull parameters, and the method of moments (Lindgren 1976) was used for the gamma distribution. Let $F_W(x)$ and $F_G(x)$ denote the Weibull and gamma distribution functions respectively with parameter values estimated from data.

Goodness of fit is based upon the empirical distribution

$$F_S(x) = \begin{cases} 0 & < x_{(1)} \\ i/n & x_{(i)} \leq x < x_{(i+1)} \quad i = 1, 2, \ldots, n \\ 1 & x_{(n)} < x \end{cases} \qquad \ldots\ldots\ (1)$$

where $x(1)$, $x(2)$, ..., $x(n)$ are the ordered data values. Two common criteria were used to judge the fit. The sum of squared deviations, i.e.,

$$SS = \Sigma(F_W(x_{(i)}) - F_S(x_{(i)}))^2$$

for the Weibull fit or

$$SS = \Sigma(F_G(x_{(i)}) - F_S(x_{(i)}))^2$$

for the gamma fit. The second measure is a Kolomogorov type (denoted K) where

$$K = \Sigma|F_W(x_{(i)}) - F_S(x_{(i)})|$$

or

$$K = \Sigma|F_G(x_{(i)}) - F_S(x_{(i)})|$$

for the Weibull or gamma distributions respectively.

According to the first measure of fit (SS), three times out of the 20 runs the gamma exhibited the better fit. In eight out of the 20 runs, the second measure (K) showed the gamma as having the better fit. This frequency of misclassification demonstrates a real possibility of selecting the wrong distribution with real data.

The log-Pearson type III distribution is the most widely used for flood frequency analysis. It has been chosen from among several candidate distributions by first estimating the parameters of each distribution for each of a large number of gaged records (Benson 1968). Then a goodness-of-fit criterion which emphasizes selected flood flows from 2 to 100 years (U.S. Water Resources Council 1976, Appendix 14) was used to select the best overall fit. Although selection of the log-Pearson type III is based upon fit in the right tail, estimation of parameters for each distribution is by standard methods which emphasizes fit in the body of the data. In certain cases, the fit in the right tail is poor. Even if the fit is good, blind application of a distribution selected on the basis of empirical fit can lead to serious error. The magnitude of this error is illustrated in the following example. The 99th percentile was computed from both the Weibull and gamma estimated distribution for each of the 20 data sets. The results are summarized in Table 1. In every case the gamma distributed percentile exceeded the true value and the Weibull estimated value. The average Weibull estimate also exceeds the true value, however the amount is within the expected sampling variation for the mean of 20 samples. Considerable overestimation bias is exhibited by the gamma distribution. This bias can be serious because overestimation can lead to a design that is too large or an estimate of the probability of failure of existing structures that is too large. Obviously, factors besides empirical fit need to be considered in selecting a distribution to fit a data set.

Table 1. Ninety-ninth percentile averages.

Data Set	True Value	Gamma Estimate	Weibull Estimate
All 20 runs	38.70	42.24	39.71
3 runs with Gamma best by SS	38.70	42.58	40.04
8 runs with Gamma best by K	38.70	42.81	40.27

EXTREME VALUE APPLICATION - HOMOGENEOUS DATA

Given the need to supplement empirical fit with theoretical considerations, the purpose of this section is to evaluate extreme value theory as a tool in identifying a distribution for annual floods. It should be understood that in all likelihood no single distribution is correct for all flood series. For example, river basins with large carry-over storage or streams which flow only intermittently may violate the assumptions of extreme value theory. In the first case, flood peaks depend on flows in the previous year; and in the second, a data set with large numbers of zero flows is not really an extreme value situation.

However, if the theory can be shown to apply in more normal situations, the hypotheses of the theory are sufficiently general to expect it to be widely applicable. In this section a theoretical distribution is selected by matching physical characteristics of stream flow with the mathematical characteristics of the various extreme value forms. Applicability is examined by trying to fit the data for selected stations with long periods of record from around the world (Table 2) used in the study of Bobee and Robitaille (1977). (See Appendix H.) The same measures of goodness-of-fit is used in order to compare these results with those obtained from the distribution of their study.

Extreme Value Distributions

Before proceeding, some basic elements of extreme value theory need to be reviewed. Extreme value random variables are defined as follows. Let x_1, x_2, ..., x_n be a sample of independent, identically distributed, continuous random variables. Let

$$Z_n = \max(x_1, x_2, \ldots, x_n) \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and

$$Y_n = \min(x_1, x_2, \ldots, x_n) \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Extreme value theory is concerned with the asymptotic distribution of sequences $(Z_n - b_n)/a_n$ and $(Y_n - b_n')/a_n'$ as n = 1,2,..., . The norming values a_n, b_n, a_n', b_n' are dictated by the theory. The interesting result of the theory is that if an asymptotic distribution exists, there are only three types for Z_n and three types for Y_n. The mathematical characteristics for the random variables x_1 which determine the resulting distribution for Z_n and Y_n are given by Gnedenko (1943). These results are difficult to use because the distribution function must be known. A less mathematical but more workable approach is suggested here.

Table 2. Selected stations exhibiting homogeneous sources.

Station	Country	River	Location	Drainage Area, Km^2	Record	Missing Years	Years of Record
bB24	Mali	Senegal	Bakel	218,000	1903-1966		64
HE60	USA	Susquehanna	Harrisburg, PA	62,400	1891-1967	1906,1922,1927 1935,1938,1951	70
IB06	India	Krishna	Vijayawada	251,355	1901-1960		60
BF40	Czech.	Decin	Elbe	51,104	1851-1968	1857,1863,1866,1873 1874,1879,1884,1898 1918,1921	108
BE38	Germany	Hofkirchen	Danube	47,495	1901-1968		68
BF19	Norway	Gloma	Langnes	40,170	1902-1968	1964	66
CF25	USSR	Neman	Smalininkai	81,200	1812-1969	1944,1945,1946	155
mE19	Canada	Hope	Fraser	203,000	1912-1970		59
JE792	Canada	Headingley	Assinibione	162,000	1914-1970		57
IF00	Canada	Medicine Hat	S.Saskatchewan	58,400	1913-1970		58
KF62	Canada	Saskatoon	S.Saskatchewan	139,500	1912-1970		59
KF53	Canada	Prince Albert	N.Saskatchewan	119,500			59
hE88a	Canada	Amos	Hurricana	3,680	1915-1969	1932,1933	53
JF50a	Canada	Slave Falls Power Plant	Winnipeg	126,000	1908-1970	1909,1911-1912,1917 1922-1926,1931,1934 1939-1942,1949,1958 1961,1962,1964,1965 1967	50

Since flood frequency analysis deals with maximum flows, only the distribution of Z_n is considered. The three possible distributions of Z_n are (Gnedenko 1943),

$$F_1(x) = \exp\left\{-\exp-\left(\frac{x-b}{c}\right)\right\} \quad -\infty<x<\infty \quad , c>0 \quad . \quad . \quad . \qquad (4)$$

$$F_2(x) = \begin{cases} 0 & x<b \\ \exp\left\{-\left(\frac{x-b}{c}\right)^{-a}\right\} & x\geq b, \quad c>0,\ a>0 \end{cases} \quad . \quad . \quad . \qquad (5)$$

$$F_3(x) = \begin{cases} 1 & \\ \exp\left\{-\left(\frac{b-x}{c}\right)^{a}\right\} & x\leq b, \quad c>0,\ a>0 \end{cases} \quad . \quad . \quad . \qquad (6)$$

Qualitative characteristics of these distributions are discussed in the next section. The assumption of independence of the x_1, x_2, ..., x_n random variables is violated in many applications. However, Watson (1952) has shown that independence is not a necessary assumption. If the randomized sequence of x_i's satisfies the assumption for all n, the theory holds.

The advantage of the theory is that once an extreme value situation is recognized one can legitimately confine the search for best fit to three extreme value distributions. The mathematical characteristics of the three distributions are very different, thus it is relatively easy to determine the correct one for a given set of data. A graphical procedure is given below for use in identifying which of the extreme value distributions should be used with a given set of data.

Determining Extreme Value Type

Distributions (4), (5), and (6) have some easily observed characteristics. The function $F_3(x)$ is limited to some maximum value b (i.e., $F_3(x) = 1$ for $x \geq b$), thus random variables which have an upper limit have extreme value form $F_3(x)$. The converse of this statement is not necessarily true, however, and variables which are not limited may also have this form (Gnedenko 1943).

The form $F_2(x)$ is referred to as a "Cauchy type" because the extreme values for the Cauchy distribution follow distribution (5). Cauchy type distributions are "heavy tailed" and seldom occur in nature. Thus, distribution (5) has limited usefulness compared with the other two types. There is, however, reference to its use in Gumbel (1954). The form $F_1(x)$ is the one most widely used and generally the only one explained in textbooks.

Three simple plots constitute the easiest method of determining which extreme value distribution is appropriate. Let $x_{(1)}$, $x_{(2)}$, ..., $x_{(n)}$ represent the ordered extreme value data for the observed maximums.

For any random variable, the expected value of its distribution function evaluated at the ith order statistic is i/(n+1) where the sample size is n (i.e., $E(F(x_{(i)})) = i/(n+1)$) (Lindgren 1976). Define $E_i = i/(n+1)$. Note that from Equation 4

$$\ln(-\ln F_1(x_{(i)})) = -x_{(i)}/c + b/c \quad . \quad . \quad . \quad . \qquad (7)$$

Note that the relationship in Equation 7 is linear in $x_{(i)}$. Substituting E_i for $F(x_{(i)})$ in Equation 7 and plotting $X_{(i)}$ vs. ln (-ln $F(x_{(i)})$) identifies data from a population with distribution function $F_1(x)$. If Equation 4 is appropriate the plot will be a straight line as illustrated in Figure 1. If the data are from any other distribution, the plot will not be a straight line.

The plot which identifies data from an $F_2(x)$ population is similar. From Equation 5 it follows that

$$\ln(-\ln F_2(x_{(i)})) = -a\ln(x-b) + a\ln c \quad . \quad . \qquad (8)$$

Thus if data are from a population with distribution $F_2(x)$, the plot of $\ln(x_{(i)} - b)$ vs. ln (-ln E_i) will be a straight line with negative slope as illustrated in Figure 2. The parameter b must be estimated before the plot can be made. Estimation of parameters is considered later.

The third plot which identifies $F_3(x)$ is motivated from Equation 6 in the same manner, i.e., the plot of ln (b - $x_{(i)}$) vs. ln (-ln E_i) is a straight line with positive slope as illustrated in Figure 3.

As discussed by Bobee and Robitaille (1977), the physical limitations of meteorological phenomena and basin characteristics which control river flow suggest that flows are bounded by an upper limit. Thus it seems that the most logical distribution for the statistical description of flood peaks is $F_3(x)$. Figure 4 verifies this choice for the Kymijoki River in Finland. It is very evident from a glance that the data are linear in this case. In less obvious cases, standard analysis techniques can be used to test for linearity (the existence of higher order polynomial effects).

In order to interpret the plot for $F_3(x)$, it is useful to examine the shape of this plot if the data were to originate from a Pearson or log Pearson type III distribution. Relative to these distributions, if floods are bounded above, the general shape of ln (b - $x_{(i)}$) plotted against ln (-ln E_i) is a curve, concave as viewed from the left. If floods are bounded below, the plot will appear as a curve convex as viewed from the left. Note that for this plot an upper bound is estimated as if the distribution were $F_3(x)$ even though it is not.

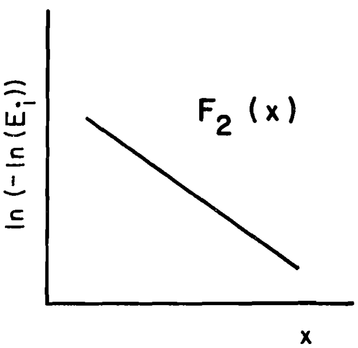

Figure 1. Straight line plot.

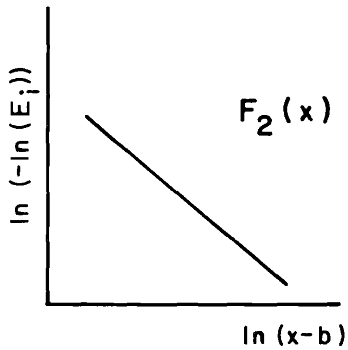

Figure 2. Straight line with negative slope.

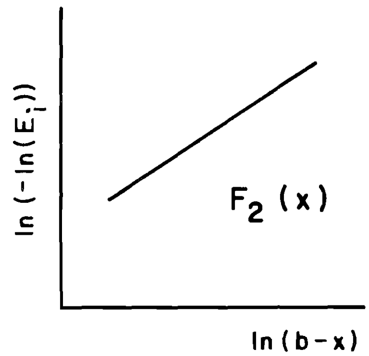

Figure 3. Straight line with positive slope.

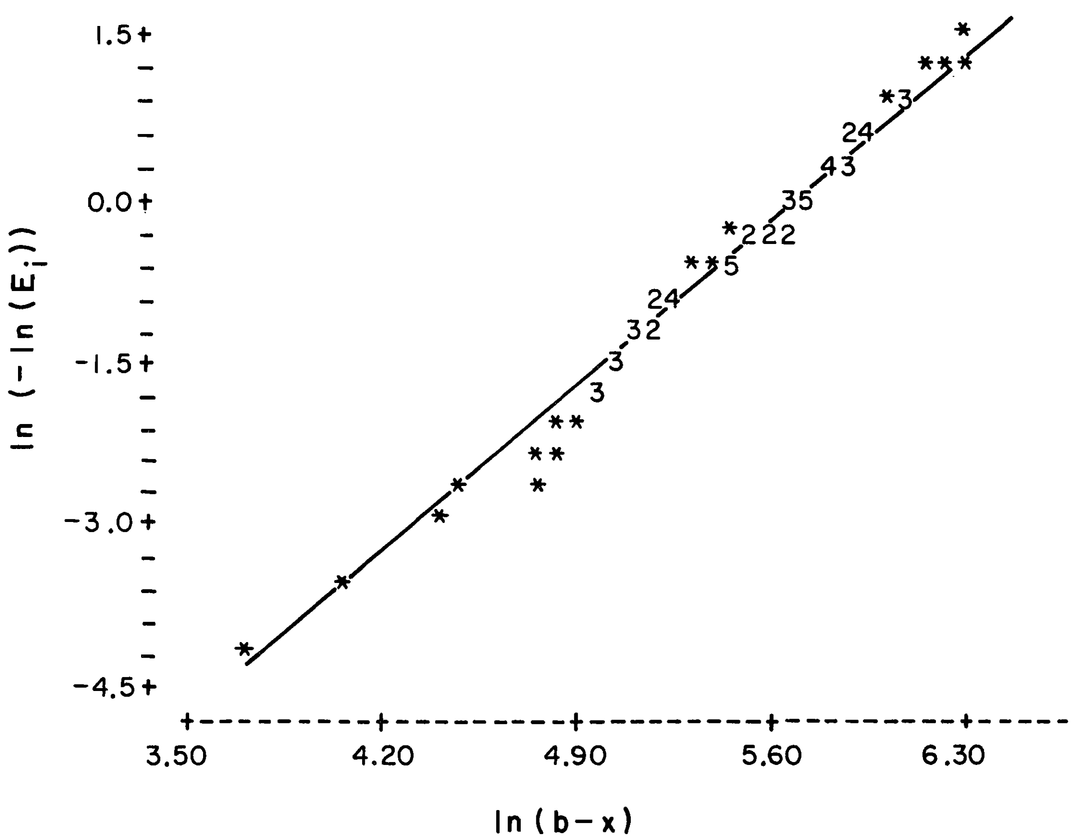

Figure 4. Verification for the Kymijoki River in Finland.

It is interesting to note that in the work of Bobee and Robitaille (1977), both the Pearson type III and log Pearson type III distributions introduce an apparent inconsistency. In some cases an upper bound for annual floods is appropriate and in others a lower bound is used. The Pearson and log Pearson distributions are not even consistent for a given data set. In some cases the Pearson distribution calls for an upper bound while the log Pearson calls for a lower bound. It seems that if an upper bound is valid due to meteorological and geographical limitations, it would be valid for all systems. The switch in boundedness is due to the inability of the Pearson and log Pearson type III distributions to accommodate both positive and negative skewness for a given bound (upper or lower).

Estimation of Parameters

Although the concept of limiting flood is reasonable, its magnitude is difficult to estimate from geographical considerations. It was found, however, that the flow estimated for a given frequency is very insensitive to the value chosen for b as long as it is relatively large. Therefore, ordinary maximum likelihood estimates of all of the parameters were used.

The distribution $F_3(x)$ is a transformed Weibull, i.e., if the $F_3(x)$ is transformed by $y = -x$ the distribution of y is Weibull with the same parameters as $F_3(x)$ (b is negative). Therefore a program available for maximum likelihood (ML) estimation of Weibull parameters (Harter and Moore 1965) was used (Appendix G). This program and other procedures described later in the report requires that the data be ordered. A FORTRAN program for this purpose is found in Appendix A.

Some difficulties were experienced in applying ML methods. In general, the computer program was expensive to run and, in addition, required several passes to find acceptable scale factors and initial values. The resulting estimates were highly dependent on these values even when the convergence criterion for the computation was met. In some cases, a better fit was obtained using a less stringent convergence measure. These problems motivated additional research not directly connected with this project.

This research resulted in a computationally more efficient method of estimation developed for all extreme value distributions (Kwan 1979). This method of estimation does not depend upon sensitive convergence criteria. These results were obtained too late to be incorporated into the comparisons made in this report. It is felt that improvement in the goodness-of-fit statistics for some of the streams reported in the next section could be obtained using the new method of estimation.

Goodness-of-fit Comparisons

The result of fitting $F_3(x)$ to the same data used by Bobee and Robitaille (1977) (Table 2) to evaluate the Pearson and log Pearson type III distributions is given in this section. Maximum likelihood estimation (with its accompanying difficulties) was used. The same goodness-of-fit statistics used by Bobee and Robitaille (1977) are used herein. These statistics are derived from three formulas for expected probabilities of order statistics referred to as the Hazen, Chegodayev, and Weibull formulas. A detailed description of the goodness-of-fit computations is given in Bobee and Robitaille (1977). Briefly the measures are based upon the relative deviations,

$$q(T) = \frac{Q(T) - D(T)}{D(T)} * 100$$

where D(T) represents the empirical (data value) for recurrence interval T, and Q(T) represents the value estimated from the fitted distribution. The recurrence intervals T = 2, 5, 10, 20, 50, and 100 were used. The average absolute deviation (i.e., $\sum_T |q(T)|/L$) is given in Table 3, and the average of the quadratic deviations (i.e., $\sum_T q(T)^2/L$) is given in Table 4. FORTRAN programs for these computations are found in Appendices D, E, and F. The goodness-of-fit values for the log Pearson type III distribution and for the distribution and method of fitting judged best by Bobee and Robitaille (1977) (Pearson type III) are also tabulated in Tables 3 and 4 for comparative purposes.

It is impossible to interpret the information on Tables 3 and 4 without viewing plots of these data sets. The plots are shown in Figures 5-18.

It can be seen that Figures 5, 10, and 17 (for stations bB24, jF50a, and BF19 respectively) have linear plots indicating an $F_3(x)$ distribution. The goodness-of-fit statistics tabulated in Tables 3 and 4 bear out this choice as the fit for $F_3(x)$ is best for the data at these three stations. The "S" shape of the plots in Figures 7, 8, 11, 12, 13 and 18 indicate that neither $F_3(x)$, Pearson type III nor log Pearson type III distributions are appropriate. These plots underscore their importance in fitting data. Whenever several distributions are fit to given data, one will always have a "best" fit. However, none of those tried may be appropriate. The plots identify these cases.

One physical explanation for a situation in which the data do not plot as a straight line is that they may not come from a single homogeneous source. The effect of non-homogeneous sources is investigated in the remaining sections of this report. The very good fits in association with the plots clearly establish extreme value theory as a viable tool for describing annual flood events.

Table 3. Mean of the absolute relative deviations.

Station	Pearson Type III			log Pearson Type III			$F_3(x)$		
	H[a]	C[a]	W[a]	H	C	W	H	C	W
bB24	1.4	1.7	2.1	1.8	1.7	2.1	1.6	1.4	1.6
hE60	3.6	4.0	4.9	3.7	3.5	4.3	7.5	5.4	5.4
IB06	3.4	2.9	3.4	3.3	3.8	4.7	7.4	7.4	8.3
BF40	3.6	4.2	4.2	3.8	4.7	4.8	7.7	7.8	8.4
BE38	3.1	2.9	2.4	2.5	2.4	2.4	2.7	2.1	3.9
BF19	3.5	4.0	4.0	3.5	4.1	4.1	3.4	3.9	4.0
CF25	2.8	2.9	3.3	3.3	3.3	3.6	7.4	6.1	6.5
mE19	2.7	2.2	3.4	2.5	2.1	3.3	3.4	2.8	3.8
jE792	7.6	5.8	6.1	6.2	5.1	4.8	6.4	6.3	6.8
iF00	2.9	4.1	5.9	4.2	5.9	7.7	15.8	17.1	15.5
kF62	4.8	4.5	4.5	4.8	5.8	5.8	10.4	11.3	11.3
kF53	6.6	4.6	6.8	6.6	4.8	8.5	13.7	11.2	14.5
hE88a	1.4	1.8	2.8	1.7	2.5	3.5	1.8	2.3	2.5
jF50a	4.4	3.6	4.4	3.8	3.4	4.2	4.2	4.4	5.4

[a]H = Hazen Formula
C = Chegodayev Formula
W = Weibull Formula

Table 4. Mean of the quadratic deviations.

Station	Pearson Type III			log Pearson Type III			$F_3(x)$		
	H[a]	C[a]	W[a]	H	C	W	H	C	W
bB24	2.9	4.1	9.4	4.3	5.1	11.2	5.0	3.4	4.6
hE60	13.4	17.6	32.3	18.9	20.8	41.3	101.0	56.9	56.9
IB06	20.4	21.2	28.2	24.0	32.1	43.6	87.7	95.8	121.1
BF40	18.0	21.9	23.7	21.9	27.7	30.9	75.7	80.1	91.4
BE38	16.2	10.2	7.0	11.0	7.1	8.7	9.6	8.1	20.9
BF19	14.5	17.7	19.7	15.7	19.6	22.2	14.0	17.2	19.3
CF25	14.2	15.2	16.0	17.6	18.4	20.1	95.1	72.45	77.2
mE19	10.7	6.6	20.7	10.6	5.8	22.7	14.2	10.7	19.8
jE792	81.4	47.8	49.5	47.6	33.1	33.7	59.4	63.6	72.9
iF00	11.4	19.2	40.9	29.2	45.1	72.8	297.0	351.0	228.9
kF62	23.9	20.7	21.7	26.0	34.5	35.8	122.7	157.2	163.6
kF53	81.3	41.3	82.0	55.6	26.8	122.8	312.0	192.4	380.4
hE88a	2.6	4.5	11.5	4.4	7.6	16.5	4.2	6.9	8.2
jF50a	31.7	13.8	21.7	21.7	13.3	22.2	22.7	24.1	37.1

[a]H = Hazen Formula
C = Chegodayev Formula
W = Weibull Formula

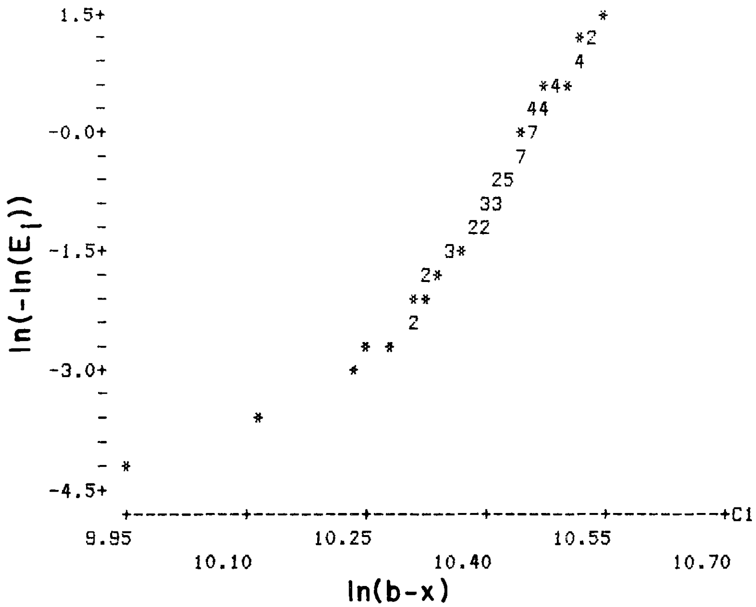

Figure 5. Station bB24--Mali River.

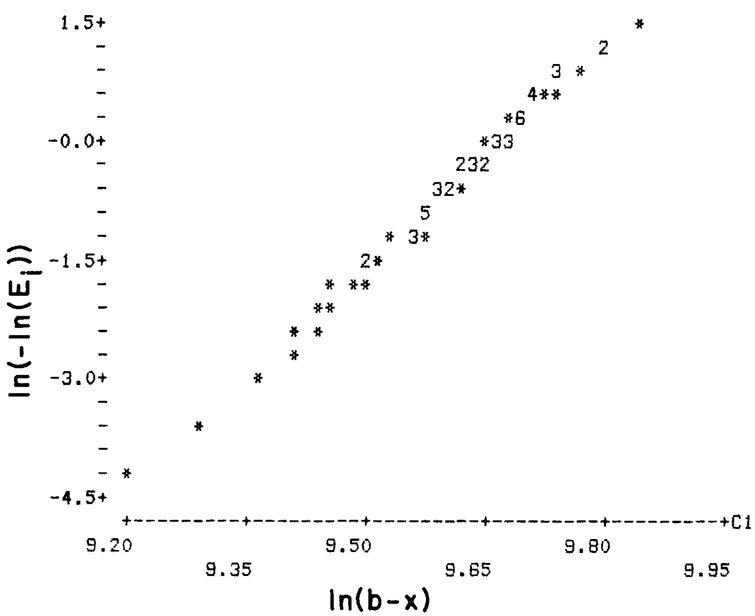

Figure 6. Station HE60--Susquehanna River.

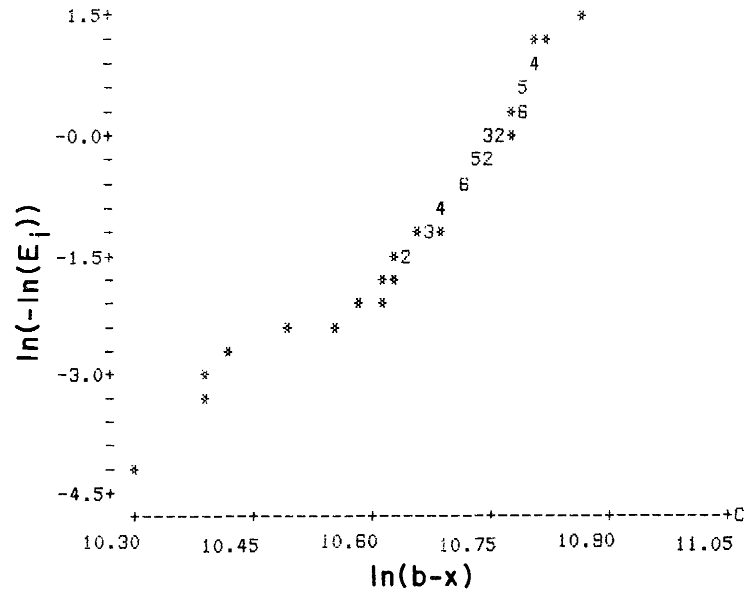

Figure 7. Station IB06--Krishna River.

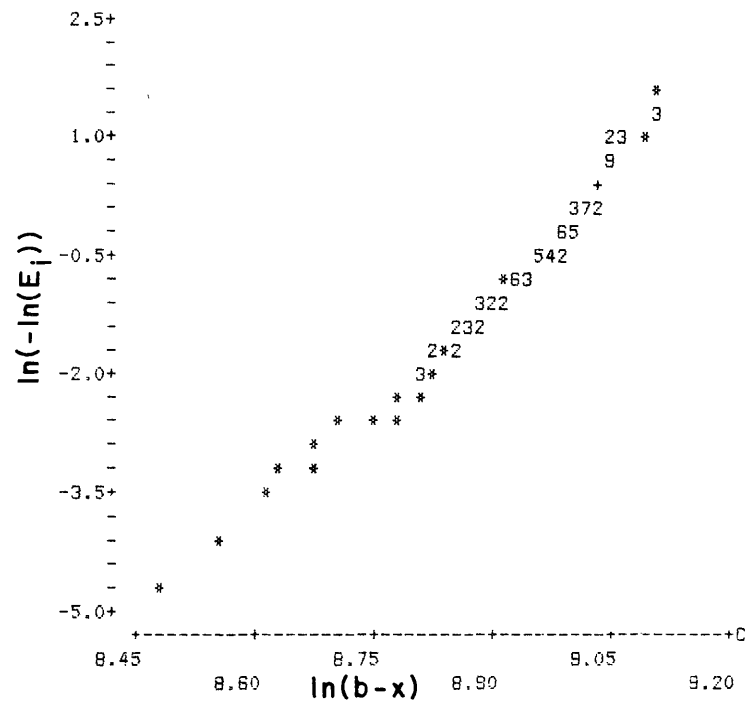

Figure 8. Station BF40--Elbe River.

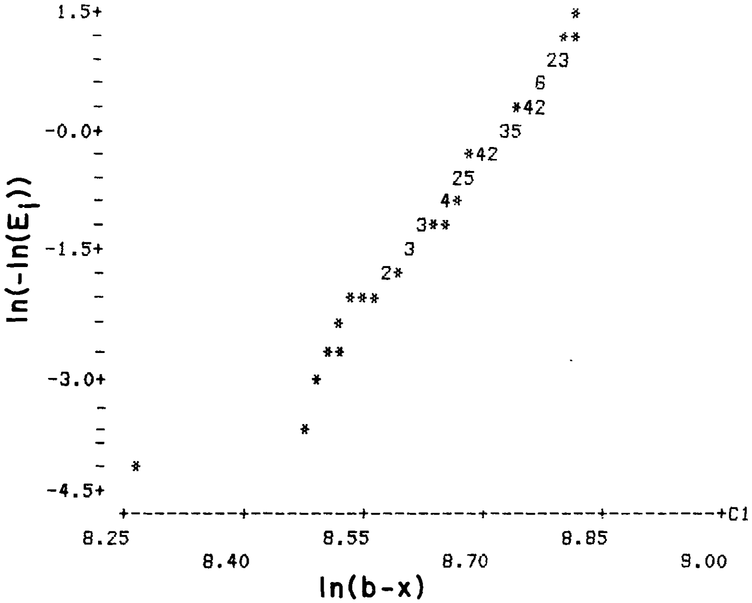

Figure 9. Station BE38--Danube River.

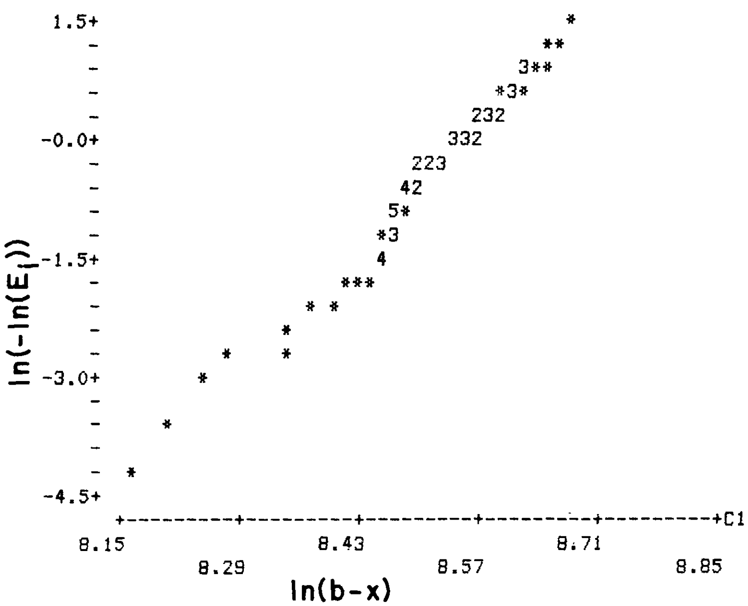

Figure 10. Station BF19--Gloma River.

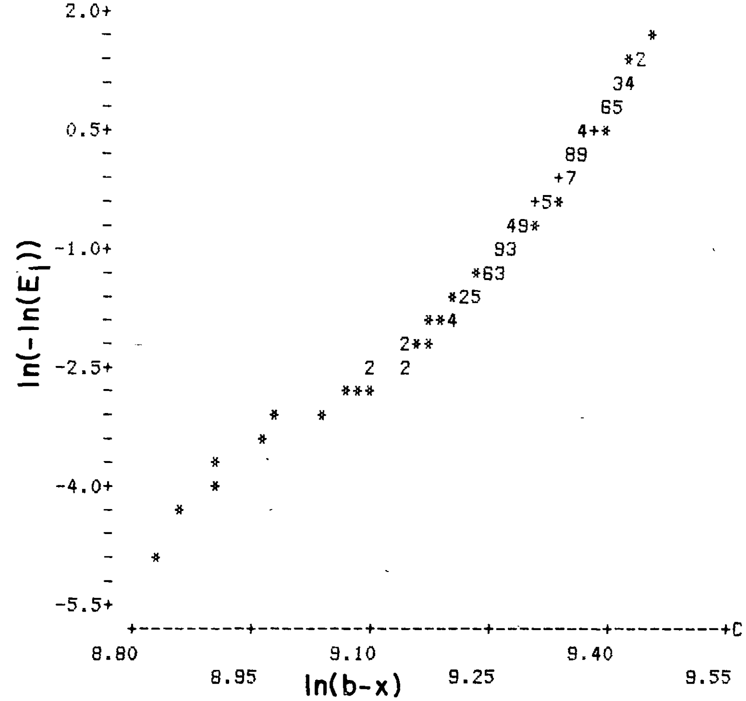

Figure 11. Station CF25--Neman River.

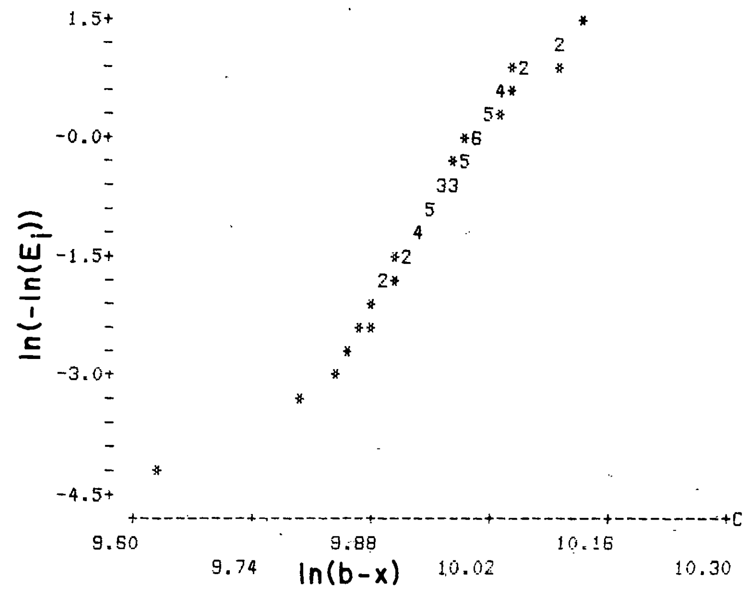

Figure 12. Station ME19--Fraser River.

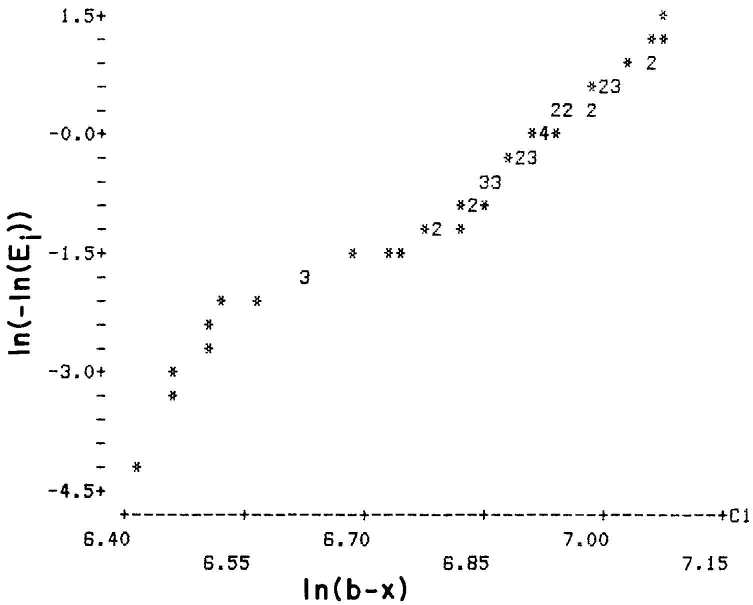

Figure 13. Station JE792--Headingly River.

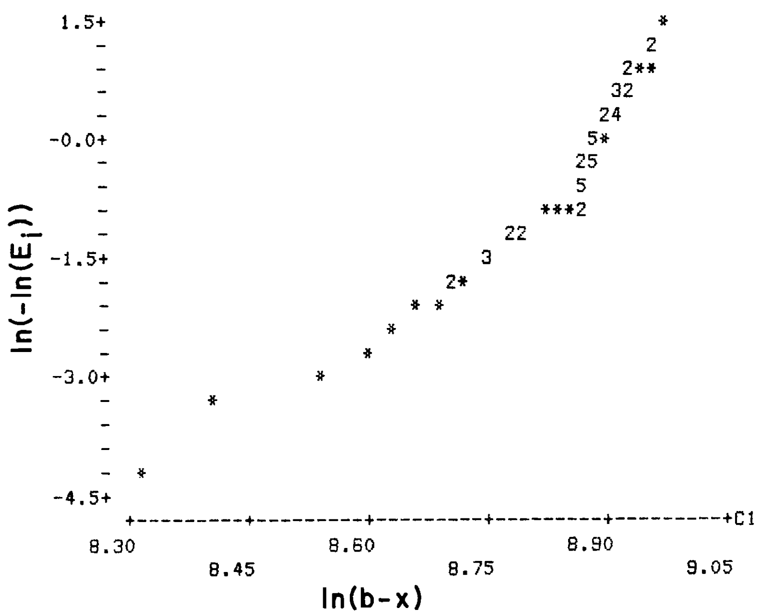

Figure 14. Station IF00--Medicine Hat River.

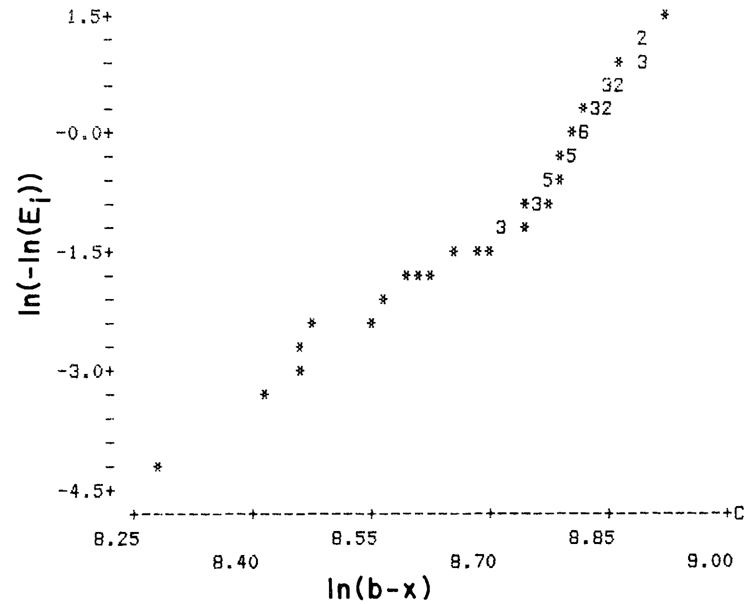

Figure 15. Station KF62--Saskatoon River.

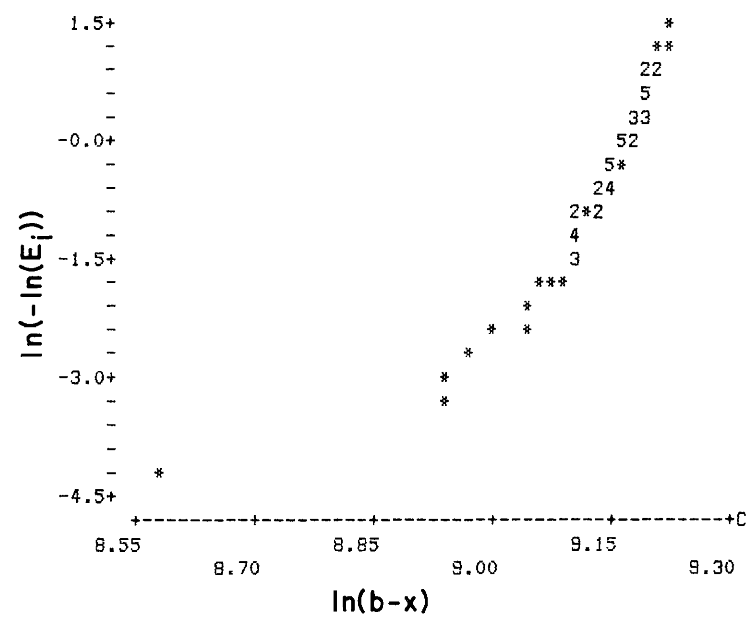

Figure 16. Station DF53--Prince Albert River.

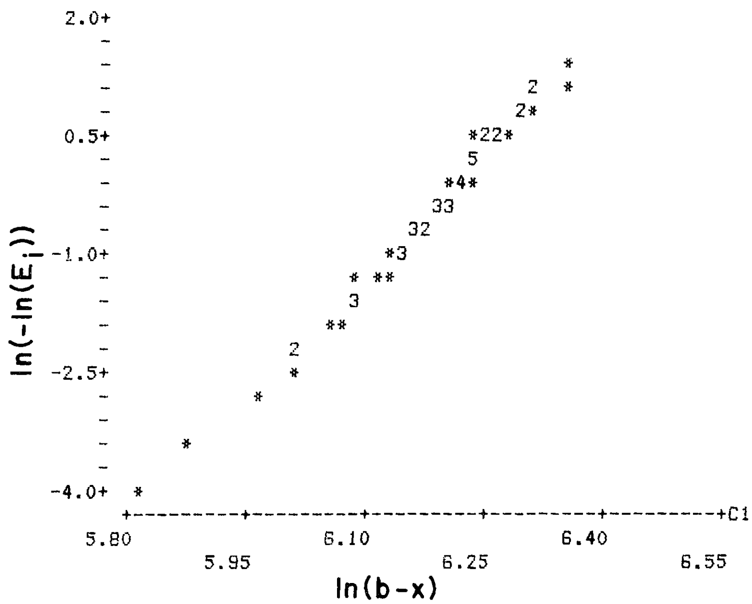

Figure 17. Station hE88a--Ames River.

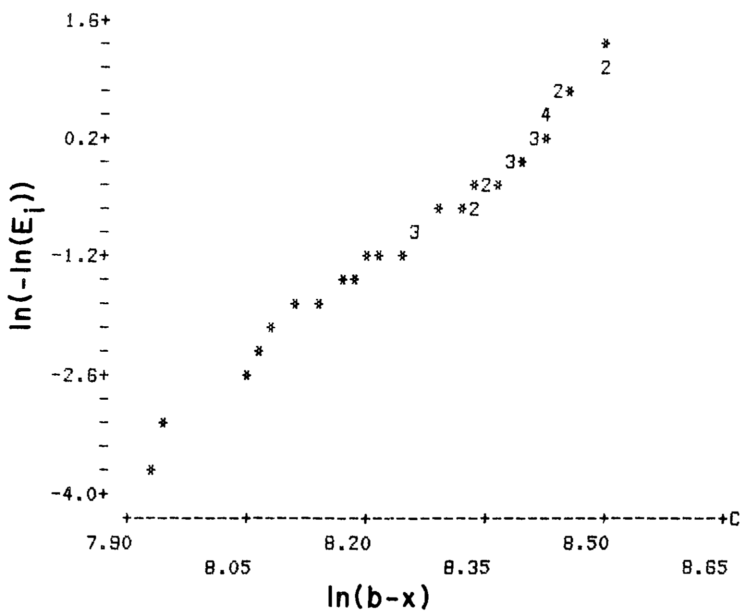

Figure 18. Station FJ50a--Slave Falls River.

EXTREME VALUE APPLICATION-NONHOMOGENEOUS DATA

Sometimes, the breaks in the slopes of the lines in plots like Figures 5 through 18 are because the data come from more than one distribution. This section of this report explores the theoretical aspects of fitting distributions to such nonhomogeneous flood data. A method of estimating the parameters of the new extreme value forms is given and the fit evaluated for several streams exhibiting nonhomogeneous sources. Identification of nonhomogeneous data by graphical methods is suggested.

Mixture Distributions in Hydrology

Prior to the observations of Ashkanasy and Weeks (1975), Potter (1958) noted the mixture of random variables in the statistical distribution of floods. He used the standard mixed distribution for the case of two components,

$$F(x) = P_1G_1(x) + P_2G_2(x) \quad . \quad . \quad . \quad . \quad . \qquad (9)$$

where $G_i(x)$, $i = 1,2$ are the distribution functions of the first and second components of the mixture respectively. The parameters P_i, $i = 1,2$ are such that $P_i > 0$, $i = 1,2$ and $P_1+P_2 = 1$. Estimation of the parameters in Equation 9 is very difficult because P_1 and P_2 must be estimated in addition to all of the parameters of both $G_1(x)$ and $G_2(x)$. Additional work by Hawkins (1972, 1974) documents other problems associated with fitting such mixed distributions.

Canfield and Borgman (1975) used reliability theory to provide a much more adequate approximating distribution. Their results have direct application to choosing a distribution of annual peak flows in hydrology in that they provide a theoretical foundation which gives primary consideration to the shape of the right tails (high flow side) of the distributions involved. Specifically, they showed the distribution of the extreme in a sequence of mixture random variables to be

$$F(x) = F_i(x)^{P_1} \, F_{i'}(x)^{P_2} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (10)$$

where the components $F_i(x)$ and $F_i'(x)$ are extreme value distributions (4), (5), or (6). Note that the parameters P_1 and P_2 can be absorbed by reparameterization so that Equation 10 can be rewritten,

$$F(x) = F_i(x)F_{i'}(x) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (11)$$

thereby reducing the number of parameters in the distribution. Because of its theoretical basis, a distribution of this form should have the correct tail characteristics. Note that the tail shape in Equation 9 is a weighted average of the tails of $G_1(x)$ and $G_2(x)$, whereas the shape of Equation 11 is a product of the tails of $F_i(x)$ and $F_i'(x)$. Even if two extreme value distributions are used in Equation 9, the tail shape is not necessarily correct.

Estimation of Parameters

The usefulness of the distributions described in the previous section depends upon 1) the availability of techniques for estimating parameter values and 2) a theoretical justification of the distributions. Theoretical justification depends on the applicability of extreme value theory as discussed above. A graphical method of determining the best parametric form of Equation 11 and of estimating the parameters is given in this section.

Graphs should always be used as a part of data analysis for annual floods. They are the easiest method for selecting from among the three extreme value types as discussed previously, and in addition they easily identify nonhomogeneous sources. Application of homogeneous distributions to nonhomogeneous river data can lead to serious blunders. The graphs should be plotted and reviewed to make sure that this is not happening.

In most applications, as discussed previously, the third extreme value distribution applies, thus the form of $F_i(x)$ and $F_i'(x)$ in Equation 11 is the same for both i and i'. However, the parameter values will be different for $F_i(x)$ and $F_i'(x)$. Thus, the graphical method used in the previous discussion on homogeneous data applies here. Correct parametric forms are identified as straight lines as noted previously. For nonhomogeneous data, two or more straight lines are found.

The data used for this part of the research were those obtained from Bobee and Robitaille and identified by them as being nonhomogeneous. (See Appendix H.) Graphs of the annual flood peaks for eleven of the rivers, plotted as illustrated by Figure 3, are shown in Figures 19 to 29. As before, $F_3(x)$ is used for $F_i(x)$ and $F_i'(x)$ (i.e., $i = 1' = 3$).

Thus

$$F(x) = \begin{cases} 1 & x>b \\ \exp\left\{-\left(\frac{b-x}{c}\right)^{a}-\left(\frac{b-x}{c'}\right)^{a'}\right\} & x\geq b,\ c>0,\ c'>0 \end{cases} \quad \ldots\ldots (12)$$

The bound parameter b was taken to be the same for both components. Numerically, b is the most difficult of the three parameters to estimate and the one to which the distribution is least sensitive.

A least squares estimation technique reported in Canfield and Borgman (1975) was improved and used to estimate the parameters of Equation 12. Let $h_{(i)}$, i = 1,2, ..., n be the ith order statistic of n annual maximum flood flows. Estimates of the parameters in Equation 12 are taken to be those values which minimize,

$$\psi=\sum_{i=1}^{n} [E(\ln F(h_{(i)})) - \ln (F(h_{(i)}))]^2 W_i \quad \ldots (13)$$

where W_i is a weight factor such that

$$W_i = \frac{\text{var } (\ln F(h_{(i)}))}{\text{var } (\ln F(h_{(i)}))} \quad \ldots\ldots (14)$$

and E(·) is the expected value operator. The variance of $\ln F(h_{(i)})$ is defined by

$$Z_i = \text{var } (\ln F(h_{(i)})) = E[\ln F(h_{(i)}) - E(\ln F(h_{(i)}))]^2 \quad \ldots\ldots\ldots (15)$$

The values of $E(\ln F(h_{(i)}))$ and var $(\ln F(h_{(i)}))$ are nonparametric and may be computed using numerical integration by the trapezoid rule.

$$E[\ln F(h_{(i)})] = \frac{n!}{(i-1)!(n-i)!}\int_0^1 \ln F(h_{(i)}) [F(h_{(i)})]^{i-1} \cdot [1-F(h_{(i)})]^{n-i} dF(h_{(i)}) \quad \ldots\ldots\ldots (16)$$

$$E[\{E[\ln F(h_{(i)})] - \ln F(h_{(i)})\}^2] = E[\{\ln F(h_{(i)})\}^2] - \{E[\ln F(h_{(i)})]\}^2$$

$$= \frac{n!}{(i-1)!(n-i)!}\int_0^1 [\ln F(h_{(i)})]^2[F(h_{(i)})]^{i-1} \cdot [1-F(h_{(i)})]^{n-i} dF(h_{(i)})$$

$$-\left\{\frac{n!}{(i-1)!(n-1)!}\int_0^1 \ln F(h_{(i)}) [F(h_{(i)})]^{i-1} \cdot [1-F(h_{(i)})]^{n-i} dF(h_{(i)})\right\}^2$$

Lindgren (1976), page 218, gives the density function of the ith order statistic and, page 113, the expectation of a function of a random variable. For convenience let,

$$EL_i = E[\ln F(h_{(i)})]$$

$$ELSQ_i = E[\{E[\ln F(h_{(i)})] - \ln F(h_{(i)})\}^2]$$

$$Y_i = b - h_{(i)}$$

$$\underline{\alpha}' = (\alpha_1, \alpha_2) = (a, a')$$

$$\underline{\theta}' = (\theta_1, \theta_2) = \left(\frac{1}{c^a}, \frac{1}{(c')^{a'}}\right)$$

From this information, Equation 13 can be rewritten as

$$\psi = \sum_{i=1}^{n} \{EL_i + \theta_1 Y_i^{\alpha_1} + \theta_2 Y_i^{\alpha_2}\}^2 W_i$$

$$= \sum_{i=1} \{\sqrt{W_i}\ EL_i + \theta_1 Y_i^{\alpha_1} \sqrt{W_i} + \theta_2 Y_i^{\alpha_2} \sqrt{W_i}\}^2 \quad \ldots\ldots\ldots\ldots (17)$$

A FORTRAN program for computation of EL_i and $ELSQ_i$ are found in Appendix B. Estimation of a, a', c and c' is accomplished by estimating $\underline{\alpha}$ and $\underline{\theta}$ and then solving for a, a', c and c' respectively.

In order to minimize Equation 17, appropriate partial derivatives of ψ are evaluated and set equal to zero.

$$\frac{\partial\psi}{\partial\theta_1} = \sum_{i=1}^{n} W_i EL_i^{\alpha_2} + \theta_1 \sum_{i=1}^{n} W_i Y_i^{2\alpha_1} + \theta_2 \sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2} = 0 \quad \ldots\ldots\ldots\ldots (18)$$

$$\frac{\partial\psi}{\partial\theta_2} = \sum_{i=1}^{n} W_i EL_i Y_i^{\alpha_2} + \theta_1 \sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2} + \theta_2 \sum_{i=1}^{n} W_i Y_i^{2\alpha_2} = 0 \quad \ldots\ldots\ldots\ldots (19)$$

$$\frac{\partial\psi}{\partial\alpha_1} = \sum_{i=1}^{n} W_i EL_i Y_i^{\alpha_1} \ln Y_i + \theta_2 \sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2} \ln Y_i + \theta_1 \sum_{i=1}^{n} W_i Y_i^{2\alpha_1} \ln Y_i = 0 \quad \ldots\ldots (20)$$

$$\frac{\partial\psi}{\partial\alpha_2} = \sum_{i=1}^{n} W_i EL_i Y_i^{\alpha_2} \ln Y_i + \theta_1 \sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2} \ln Y_i + \theta_2 \sum_{i=1}^{n} W_i Y_i^{2\alpha_2} \ln Y_i = 0 \quad \ldots\ldots (21)$$

Solving Equation 18 for θ_2 yields,

$$\theta_2 = \frac{-\sum_{i=1}^{n} W_i EL_i Y_i^{\alpha_1} - \theta_1 \sum_{i=1}^{n} W_i Y_i^{2\alpha_1}}{\sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2}} \quad \ldots (22)$$

Substituting for θ_2 in Equation 19 and solving for θ_1 gives

$$\theta_1 = \left(\sum_{i=1}^{n} W_i EL_i Y_i^{\alpha_1}\right)\left(\sum_{i=1}^{n} W_i Y_i^{2\alpha_2}\right) - \left(\sum_{i=1}^{n} W_i EL_i Y_i^{\alpha_2}\right)\left(\sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2}\right) \Bigg/ \left(\sum_{i=1}^{n} W_i Y_i^{\alpha_1+\alpha_2}\right)^2 - \left(\sum_{i=1}^{n} W_i Y_i^{2\alpha_1}\right)\left(\sum_{i=1}^{n} W_i Y_i^{2\alpha_2}\right). \quad . \quad . \quad . \quad . \quad (23)$$

The result of Equation 23 is substituted into Equation 22 to yield equations for both θ_1 and θ_2 which involve the parameters α_1 and α_2 as the only unknowns. These equations are substituted for θ_1 and θ_2 in Equations 20 and 21 giving two equations in two unknowns ... α_1 and α_2. This system of equations can be solved numerically using the IMSL (1977) library subroutine ZSYSTM. Given this solution as $\underline{\alpha}$, the estimate $\hat{\underline{\theta}}$ of $\underline{\theta}$ is computed from Equations 21 and 23. Initial values of α_1 and α_2 are required in ZSYSTM. These are obtained as the slopes of the lines observed in the graph (e.g. see Figures 19 through 28). Appendix C contains FORTRAN programs for these estimates.

A Burroughs 6700 computer was used to solve for $\hat{\underline{\alpha}}$. Since the Burroughs or any other computer system is finite, a scaling factor was found to be a computational necessity, i.e., Equation 17 becomes

$$\psi = \sum_{i=1}^{n} \left\{ \sqrt{W_i}\, EL_i + (sf)^{\alpha_1} \theta_1 \left(\frac{Y_i}{sf}\right)^{\alpha_1} \sqrt{W_i} + (sf)^{\alpha_2} \theta_2 \left(\frac{Y_i}{sf}\right)^{\alpha_2} \sqrt{W_i} \right\}^2 . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

For convenience θ_1 and θ_2 are redefined so that Equation 24 may be written

$$\psi = \sum_{i=1}^{n} \left\{ \sqrt{W_i}\, EL_i + \theta_1^* \left(\frac{Y_i}{sf}\right)^{\alpha_1} \sqrt{W_i} + \theta_2^* \left(\frac{Y_i}{sf}\right)^{\alpha_2} \sqrt{W_i} \right\}^2 . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

where $\theta_i^* = (sf)^{\alpha_i} \theta_i$, $i = 1,2$.

For 8 of the 11 data sets used in this study, an adequate scale factor was the difference between the specified maximum flood and the first order statistic or smallest of the maximum yearly floods:

$$sf = b - h_{(1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

The other three data sets required manipulation of the scale factor to insure that no numbers got too large or too close to zero for the computer to handle. Of course, larger and more powerful computer facilities would lessen the importance of the scale factor.

The rivers for which data were obtained are shown in Table 5. Estimates of the parameters for each river are shown in Tables 6 and 7. It was found that the value of ψ in Equation 25 was very insensitive to b for large values of b. Therefore in order to conserve computer time, b was estimated by using a few passes to arrive at an "approximate" estimate. This procedure could be automated so that no hand preparation is necessary and slightly better estimates could be obtained. However, very little improvement is expected.

Goodness-of-fit Nonhomogeneous Data

The same goodness-of-fit statistics as described previously and used by Bobee and Robitaille were used for these data. Since the data (empirical) values of river flows for the selected return periods were not available for these rivers in Bobee and Robitaille's (1977) work, they are shown here in Tables 8, 9, and 10.

Table 5. Selected stations exhibiting nonhomogeneity in source.

No.	Station	Country	River	Location	Drainage Area, Km^2	Record	Missing Years	Years of Record
1	hE1833	Canada	Saguenay	Isle-Maligne	73,000	1913-1970		58
2	aB36	Mali	Niger	Dire	340,000	1924-1968		43
3	aB72	Mali	Niger	Koulikoro	120,000	1907-1968		62
4	aE85	USA	Penobscot	W. Enfield	17,090	1902-1967	1913,1928,1944 1951,1960,1964	60
5	CG60	Finland	Kymijoki	Pernoo	36,535	1900-1968		69
6	cG81	Finland	Vuoksi	Imatra	61,280	1847-1968		122
7	BF42	Poland	Oder	Gozdowice	109,365	1901-1968	1945	67
8	CF28	Sweden	Vanerngota	Vanesborg	46,830	1807-1968		162
9	DF09	USSR	Neva	Novosaratovka	281,000	1859-1969	1942	90
10	jE9955	Canada	Assiniboine	Brandon	92,000	1902-1970		65
11	JE791	Canada	Red	Emerson	104,000	1913-1970		58

The associated river heights (Q(T)) as estimated by Equation 12 using the respective parameters in Table 6 are shown in Table 10. The goodness-of-fit statistics are tabulated in Table 12.

It is instructive to view the plots of these rivers. Shown in Figures 19 to 29 are the plots for each river. The C_1 axis is $\ln(b-X_{(i)})$ and the C_L axis is $\ln(-\ln(i/(n+1)))$. The maximum likelihood estimated value of b has been used.

The Saguenay River (Figure 19) manifests a straight line plot and may have nearly homogeneous sources, although the two largest floods could be from another source. The Niger River, location Dire (Figure 20) and location Koulikoro (Figure 21), exhibits two sharply different components. The Penobscot River (Figure 22) appears to have homogeneous sources with close to a straight line plot. Figure 23 does not exhibit a clear indication of two sources, although there seems a tendency toward two straight lines. Its estimated parameters indicate likewise, a = 2.30 and a' = 2.30 with c = 388.32 an c' = 410.86--very close to identical components. The Vuoksi River (Figure 24)

Table 6. Maximum flood flow b (in m^3/S), scale factor sf, and parameters estimated from Equation 25.

No.	b	sf	α_1	α_2	θ_1	θ_2
1	25000	22630	16.33	8.67	4.26	0.001
2	3000	1053	3.53	719.68	3.74	0.52
3	21000	17354	14.05	14.05	0.14	6.06
4	18000	17179	0.91	27.57	0.005	2.92
5	700	562	2.30	2.30	2.34	2.06
6	2500	2167	14.55	2398.16	9.82	4.90
7	6000	5293	22.89	6.69	2.06	1.40
8	1300	1047	6.55	6.59	0.25	11.32
9	6000	4000	4.89	6.46	0.02	8.56
10	670	347.9	8.98	1.03	0.013	0.14
11	3100	1200	9.31	0.87	5.6 E-4	0.04

Table 7. Parameter estimates of a, a', c, and c' for each station.

No.	b	a	a'	c	c'
1	25000	16.33	8.67	20706.33	49696.00
2	3000	3.53	7.19	724.52	1053.96
3	21000	14.05	14.05	19932.25	15265.06
4	18000	0.91	27.57	5605113.20	16523.166
5	700	2.30	2.30	388.32	410.86
6	2500	14.55	2989.16	1852.06	2165.81
7	6000	22.89	6.69	5128.49	5033.41
8	1300	6.55	6.59	1297.56	724.48
9	6000	4.89	6.46	8904.89	2869.71
10	670	8.98	1.03	563.86	2303.15
11	3100	9.31	0.87	2683.34	49293.56

Table 8. Data values D(T) (in m^3/S) as interpolated between adjacent observations by the Chegodayev method.

	T in Years					
No.	2	5	10	20	50	100
1	4655	6125	6766	7811	9166	a
2	2335	2562	2641	2664	2677	a
3	6250	7066	7670	9065	9590	a
4	1738	2342	2342	3124	3929	a
5	454	545	578	614	648	a
6	703	794	881	933	1139	1157
7	1350	1875	2418	2759	3474	a
8	627	726	773	809	927	945
9	3300	3762	4000	4118	4500	4560
10	154	252	423	509	622	a
11	540	836	1283	1532	2300	a

[a]Beyond the range of the data.

Table 9. Data values D(T) (in m^3/S) as interpolated between adjacent observations by the Hazen method.

	T in Years					
No.	2	5	10	20	50	100
1	4655	6111	6761	7714	9128	9244
2	2335	2561	2640	2661	2677	a
3	6250	7041	7649	8964	9552	9676
4	1738	2339	2650	3081	3777	4251
5	454	545	577	614	646	655
6	703	794	880	931	1138	1153
7	1350	1867	2412	2700	3401	3655
8	627	726	733	806	927	937
9	3300	3750	4000	4100	4500	4540
10	154	251	422	498	613	644
11	540	835	1253	1518	2149	2607

[a]Beyond the range of the data.

Table 10. Data values D(T) (in m^3/S) as interpolated between adjacent observations by the Weibull method.

	T in Years					
No.	2	5	10	20	50	100
1	4655	6170	6775	7987	9224	a
2	2335	2563	2643	2670	a	a
3	6250	7103	7701	9216	9648	a
4	1738	2346	2673	3188	4156	a
5	454	546	579	615	652	a
6	703	794	881	935	1142	1164
7	1350	1888	2426	2848	3583	a
8	627	727	773	814	927	973
9	3300	3780	4000	4145	4500	4589
10	154	255	425	524	636	a
11	540	841	1310	1567	2528	a

[a]Beyond the range of the data.

has two or possibly three nonhomogeneous sources. The Oder River (Figure 25) has two components, however the definition is not sharp. The Vanerngota River (Figure 26) has well defined components and the Neva River (Figure 27) appears to be homogeneous. The Assiniboine River (Figure 28) and the Red River (Figure 29) have sharply defined components.

The goodness-of-fit for the first ten stations is excellent. The fit for the Red River is not as good although it does fit well the Weibull method of observed flood discharges. Perhaps the largest maximum yearly flood is an outlier (see Figure 29) as it is much larger than any other flood on record. Alternatively, it might be the only observation from a particular source population. It is impossible to achieve a good estimate of the parameters of a population with only one observation.

Table 11. Computed flood flows Q(T) (in m^3/S) for selected return periods.

	T in Years					
No.	2	5	10	20	50	100
1	4754	6112	6961	7740	8699	9382
2	2347	2526	2617	2688	2760	2803
3	6153	7303	8015	8664	9455	10015
4	1699	2364	2796	3212	3779	4278
5	448	546	589	619	646	660
6	694	829	913	990	1084	1150
7	1351	1988	2407	2772	3192	3470
8	617	725	787	840	901	941
9	3290	3727	3976	4190	4433	4594
10	151	262	413	542	618	644
11	552	902	1174	1622	2546	2852

Table 12. Goodness-of-fit statistics.

	Mean of the Absolute Deviations			Mean of the Quadratic Deviations		
No.	Hazen	Chegodayev	Weibull	Hazen	Chegodayev	Weibull
1	1.9	2.2	2.9	6.3	7.9	11.0
2	1.4	1.4	0.9	2.7	2.7	1.0
3	3.0	3.0	3.3	10.6	11.1	13.4
4	2.3	3.0	3.5	9.2	11.0	22.0
5	0.9	0.9	0.9	1.2	1.2	1.2
6	3.4	3.5	3.6	16.1	16.0	16.0
7	3.4	3.0	3.9	18.8	20.5	31.0
8	1.8	1.8	2.1	5.3	4.7	5.8
9	1.1	1.0	0.8	1.6	1.2	1.0
10	3.1	3.1	2.8	17.9	13.6	7.8
11	8.6	7.1	4.8	97.9	57.9	35.8

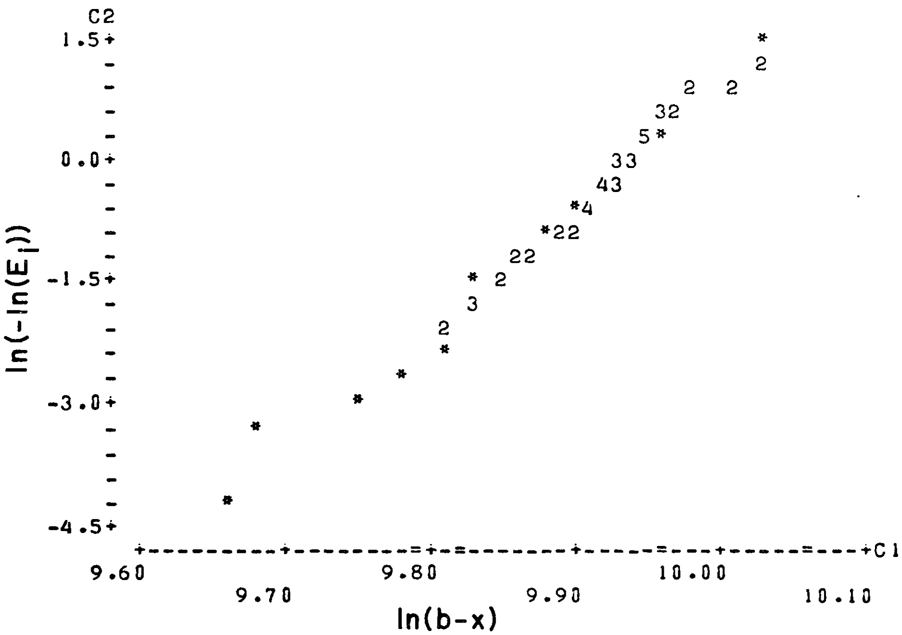

Figure 19. Station No. 1--Saguenay River.

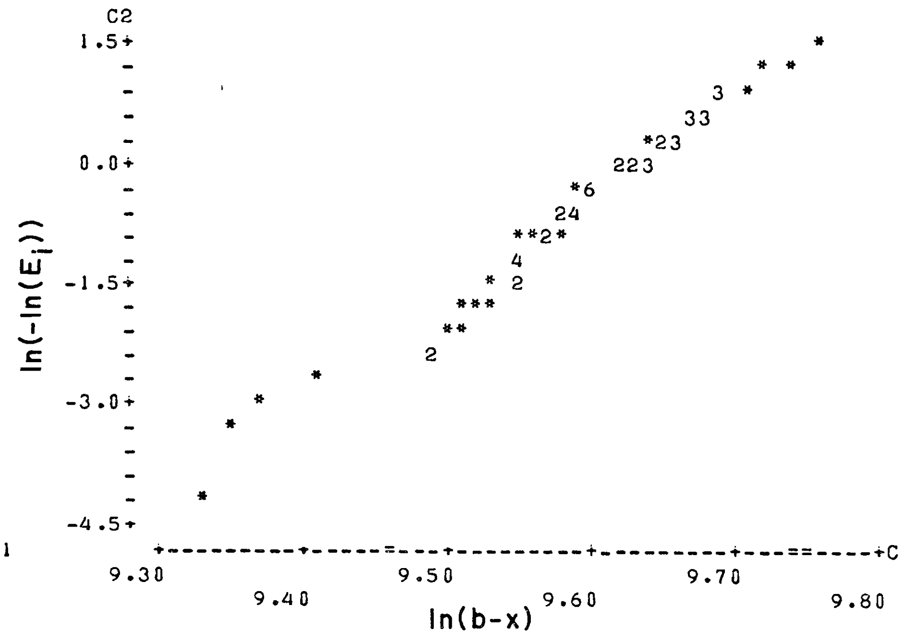

Figure 21. Station No. 3--Niger River, Koulikoro.

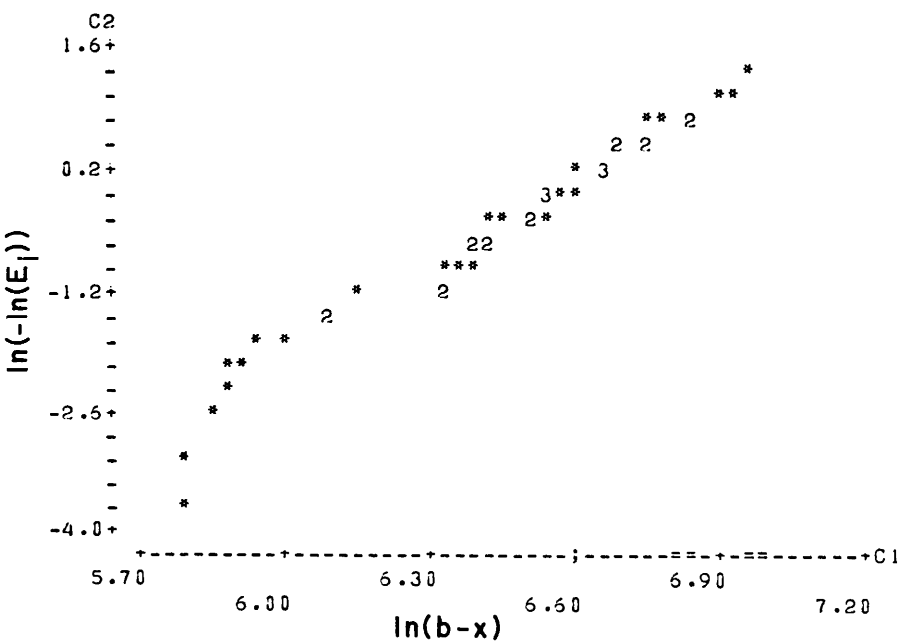

Figure 20. Station No. 2--Niger River, Dire.

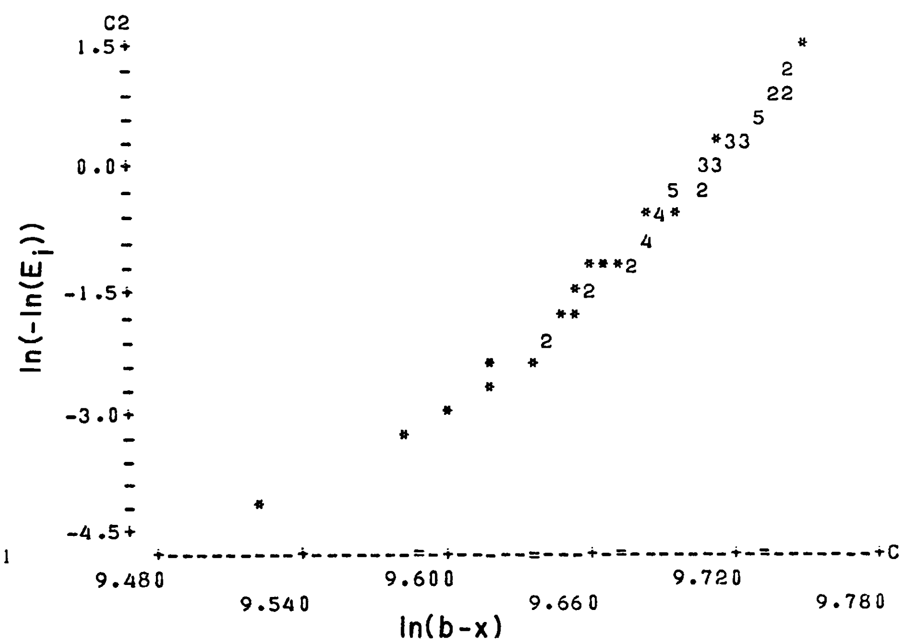

Figure 22. Station No. 4--Penobscot River.

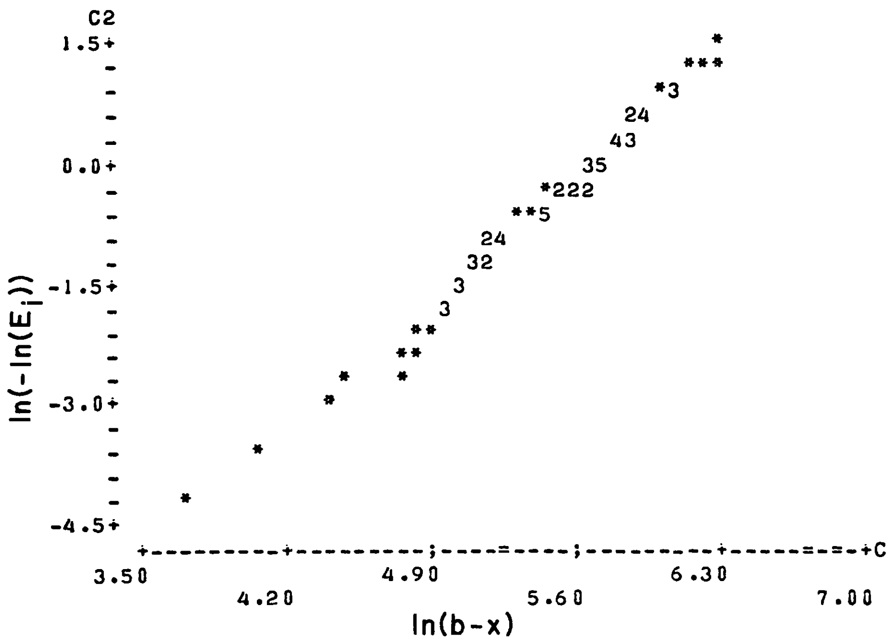

Figure 23. Station No. 5--Kymijoki River.

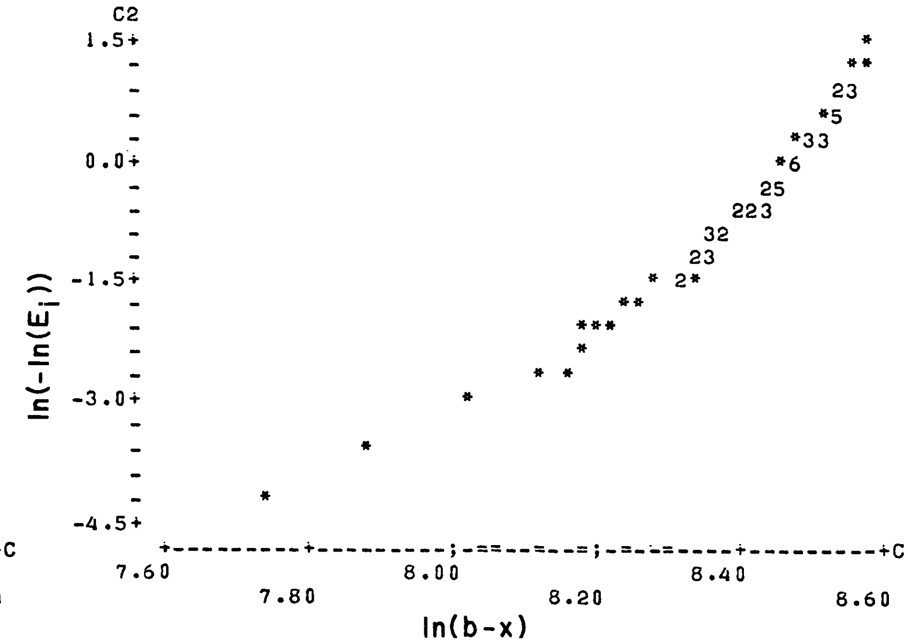

Figure 25. Station No. 7--Oder River.

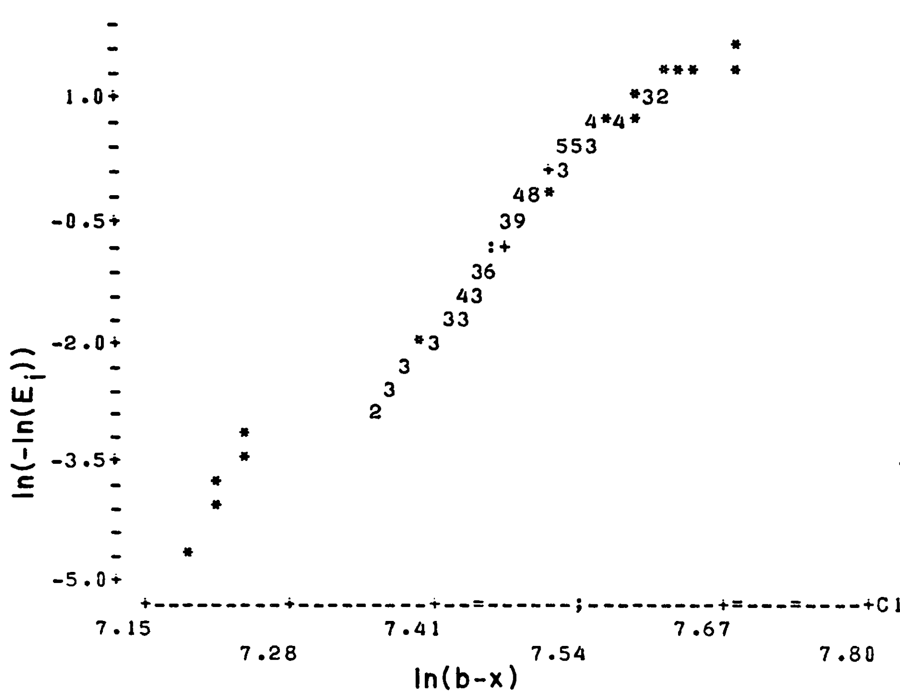

Figure 24. Station No. 6--Vuoksi River.

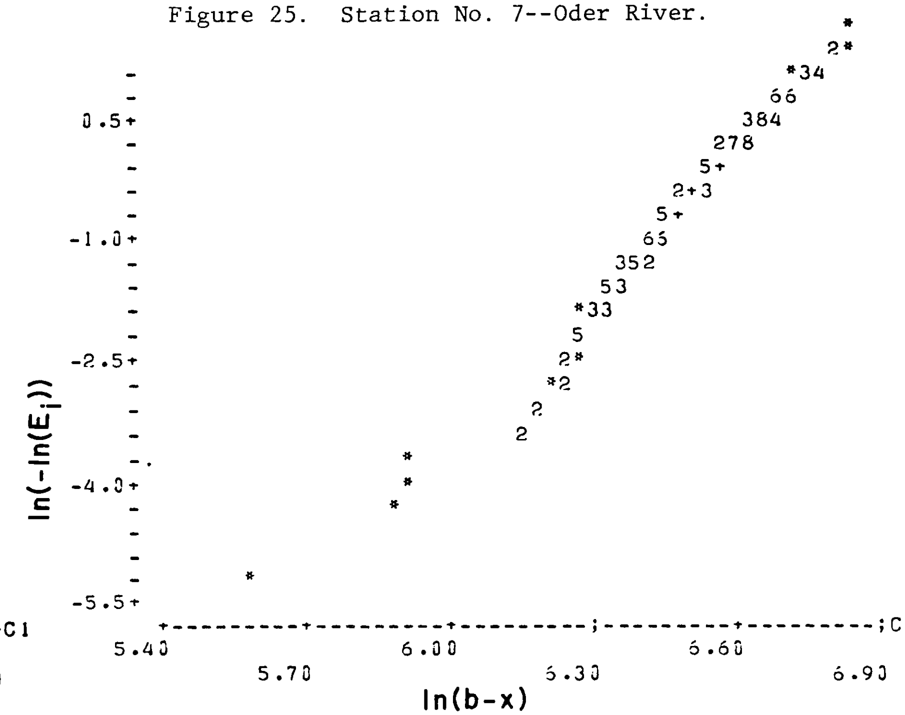

Figure 26. Station No. 8--Vanerngota River.

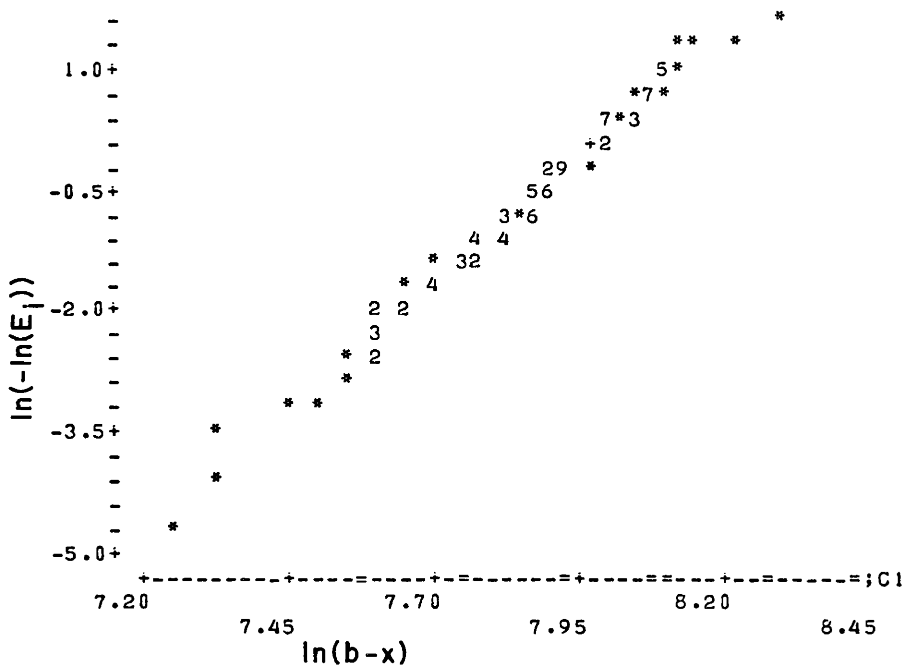

Figure 27. Station No. 9--Neva River.

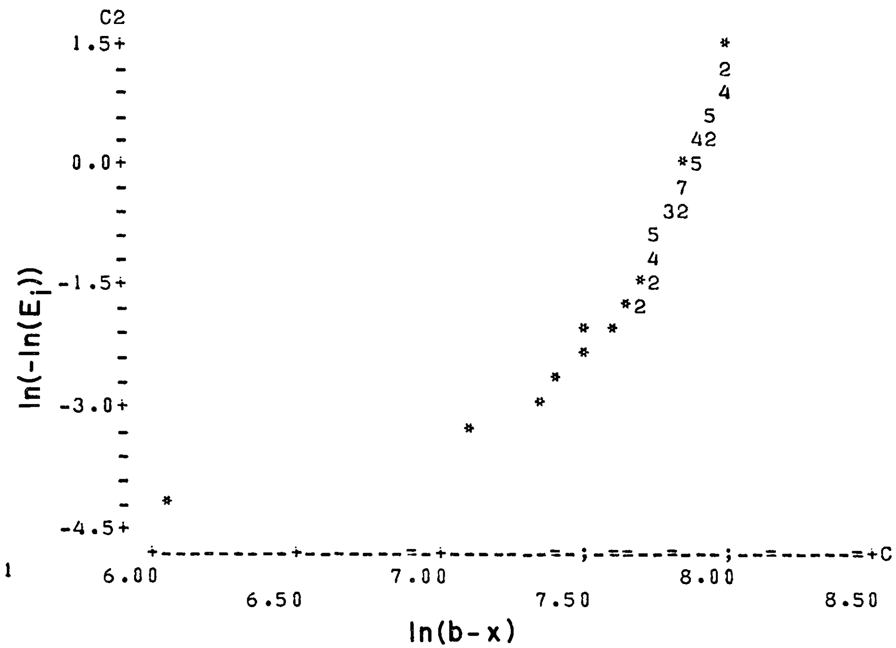

Figure 29. Station No. 11--Red River.

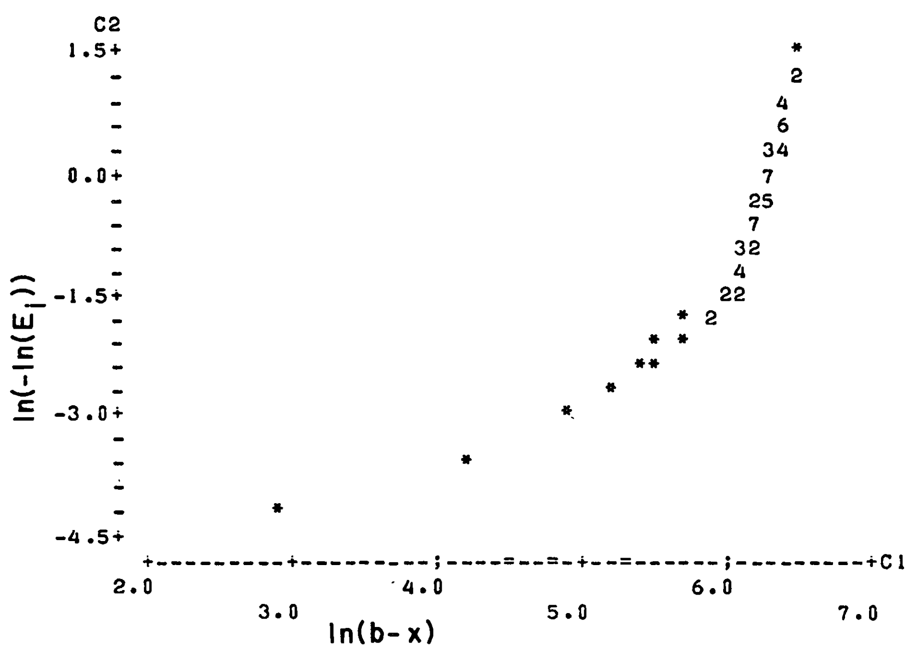

Figure 28. Station No. 10--Assiniboine River.

FLOOD FREQUENCY ANALYSIS PROCEDURE

The following steps are offered as guidelines for flood frequency analysis based on extreme-value theory as presented in this report.

1. Select a value for b in the order of two or three times the magnitude of the largest flood of record and plot the data in the form of Figure 3.

2. If the plot in Step 1 is linear, estimate parameters a, b, and c (Equation 6) and apply the results for estimating flood frequency.

3. If the plot in Step 1 is curved, some other distribution is probably more applicable, and alternatives should be considered.

4. If the plot in Step 1 exhibits a break, estimate parameters a, a', b, c, and c' (Equation 12). This is done by substituting Equations 22 and 23 in Equations 20 and 21 and solving for α_1 and α_2, estimating θ_1 and θ_2 from Equations 22 and 23, using these four values to estimate a, a', c, and c'. Computer programming lists are presented.

CONCLUSIONS

The original objective of this research was to develop and evaluate an extension of extreme value theory for application to estimating flood frequency relationships for river flows drawn from nonhomogeneous populations. Before doing so, applications to homogeneous data were considered, and a functional form that limits flows to a maximum value was found preferable to the widely used Gumbel form. A relationship was then derived for fitting data mixing two distributions. The goodness-of-fit statistics indicate excellent fit for these mixture distributions (except when one of the sources has very few observed values).

The mixture distribution, however, has five parameters and therefore should be capable of fitting a wide variety of data sets. The real justification for its application lies in its basis in extreme value theory. It was demonstrated that extreme-value distributions provide excellent fit for many river systems. The method of estimation (maximum likelihood) had some inherent difficulties which may have produced some of the poor fits. More efficient estimation methods are now available and should be tested.

Finally, extreme-value theory may not apply to all river systems. A large carry-over storage may, for example, violate the hypothesis of the theory. However, the results of this study indicate that the theory does apply to many systems.

LITERATURE CITED

Ashkanasy, N. M., and W. D. Weeks. 1975. Flood frequency distribution in a catchment subject to two storm rainfall producin mechanisms. Hydrology Symposium, Armidale, NSW, Australia, May 18-21, Institute of Eng., Sydney, Australia, p. 153-157.

Benson, M. A. 1968. Uniform flood-frequency estimating methods for federal agencies. Water Resour., 4(5):891-908.

Bobee, B. B., and R. Robitaille. 1977. The use of the Pearson type 3 and log Pearson type 3 distributions revisited. Water Resources Research, 13(2):427-443.

Canfield, R. V., and L. E. Borgman. 1975. Some distributions of time to failure for reliability applications. Technometrics, 17(2):263-268.

Gnedenko, B. V. 1943. Sur la Distribution Limite du Terms Maximum d'une Serie Aleatorie. Annals of Mathematics, 44:423-453.

Gumbel, E. J. 1958. Statistics of extremes. Columbia University Press, New York. 375 p.

Hager, H. W., L. J. Bain, and C. E. Antle. 1971. Reliability estimation for the generalized Gamma distribution and robustness of the Weibull model. Technometrics, 13(3):547-558.

Harter, H. L., and A. H. Moore. 1965. Maximum likelihood estimation of Gamma and Weibull populations from complete and censored samples. Technometrics, 7(6):639-643.

Hawkins, R. H. 1972. A note on multiple solutions to the mixed distribution problem. Technometrics, 14(4):973-976.

Hawkins, R. H. 1974. A note on mixed distributions in hydrology. Proceedings of Symposium on Statistical Hydrology, USDA, Ag. Res. Service, Misc. Pub. No. 1275, USGPO, Washington, D.C., p. 336-344.

IMSL. 1977. International Mathematical and Statistical Libraries, Inc., Library 3, Edition 6, Vol. 2, Houston, Texas.

Kwan, P. K. 1979. Weighted least squares estimation of extreme value distribution. M.S. Thesis, Department of Applied Statistics and Computer Science, Utah State University, Logan, Utah, 84322.

Lindgren, B. W. 1976. Statistical theory. Macmillan Publishing Co., Inc., New York.

Potter, W. D. 1958. Upper and lower frequency curves for peak rates of runoff. Trans. actions of the American Geophysical Union, 39:100-105.

U.S. Water Resources Council. 1976. Guidelines for determining flood flow frequency. Bulletin No. 17 of the Hydrology Committee, Washington, D.C.

Watson, G. S. 1952. Extreme value theory for m dependent stationary stochastic processes. Annals of Mathematical Statistics, 25:798-803.

White, J. S. 1969. The moments of log-Weibull order statistics. Technometrics, 11(2):373-385.

APPENDICES

Appendix A

Program ORDER

```
C     THIS PROGRAM READS THE YEARLY MAXIMUM FLOOD DATA OF A RIVER,
C     ORDERS THIS DATA INTO ASCENDING ORDER, AND THEN STORES THE
C     DATA ON DISK FOR FUTURE ANALYSIS.  NECESSARY INPUT IS THE
C     NUMBER OF YEARS OF THE RECORD AND THE ACTUAL DATA.  M IS THE
C     NUMBER OF YEARS OF DATA RECORD.  X IS AN ARRAY FOR THE DATA
C     ITSELF.
FILE  1(KIND=DISK,TITLE="SAGUENAY/DATA")

      DIMENSION X(200)
C     M, THE NUMBER OF YEARS OF DATA IS READ.
      READ(5,/)M

C     THE DATA IS READ FREE FORMAT AND STORED IN ARRAY X.
      READ(5,/)(X(I),I=1,M)
C     THE DATA IS  ORDERED IN ASCENDING ORDER, THUS X(1) IS THE
C     SMALLEST AND X(M) IS THE LARGEST MAXIMUM YEARLY FLOOD.
      NESTED=M
      L=NESTED-1
      DO 20 J=1,L
      NESTED=NESTED-1
      DO 20 I=1,NESTED
      IF(X(I)-X(I+1))20,20,30
30    SAVE=X(I)
      X(I)=X(I+1)
      X(I+1)=SAVE
20    CONTINUE
      WRITE(6,100)M

100   FORMAT(1X,' THE NUMBER OF YEARS OF RECORD=',I15,/////)
      WRITE(6,200)
200   FORMAT(1X,' THE ORDERED MAXIMUM YEARLY FLOODS',///)
      WRITE(1,101)(X(I),I=1,M)
      WRITE(6,120)(X(I),I=1,M)
120   FORMAT(1X,5F10.1,/)
101   FORMAT(1X,F12.2)
C     ORDERED DATA IS SAVED ON DISK.
      LOCK 1
      STOP
      END
```

Appendix B

Program INTEGRATE

```
C       THIS PROGRAM CALCULATES THE EL(I), ELSQ(I), AND W(I) BY
C       NUMERICAL INTEGRATION WITH THE TRAPEZOID RULE.  M, THE
C       NUMBER OF YEARS OF DATA, IS THE ONLY REQUIRED INPUT.
C       C IS THE STEP SIZE.
FILE    2(KIND=DISK,TITLE="SAGUENAY/EL")
FILE    3(KIND=DISK,TITLE="SAGUENAY/W")

        IMPLICIT DOUBLE PRECISION (A-H,O-Z)
        DIMENSION EL(200),ELSQ(200),W(200)
        C=0.01
C       M, THE NUMBER OF YEARS OF DATA, IS READ.
        READ(5,/)M

        D=M
        WRITE(6,110)

110     FORMAT(1X,16X,'EL(I)',25X,'ELSQ(I)',25X,'W(I)',///)
        DO 1 I=1,M
        EL(I)=0.
        ELSQ(I)=0.
        IF(I.EQ.1)GO TO 20
        GO TO 13
11      T1A=DLOG(F1)
        T1ASQ=(DLOG(F1))**2.
        T2A=F1**(I-1)
        T3A=(1.-F1)**(M-I)
        A=T1A*T2A*T3A
        ASQ=T1ASQ*T2A*T3A
        GO TO 14
13      A=0.
        ASQ=0.
        F2=C
14      T1B=DLOG(F2)
        T1BSQ=(DLOG(F2))**2.
        T2B=F2**(I-1)
        T3B=(1.-F2)**(M-I)
        B=T1B*T2B*T3B
        BSQ=T1BSQ*T2B*T3B
        EL(I)=(A+B)*C/2.
        ELSQ(I)=(ASQ+BSQ)*C/2.
        GO TO 10
20      F1=C/4.
        F2=C
        GO TO 11
10      A=B
        ASQ=BSQ
        F2=F2+C
        IF(F2.GT.1)GO TO 15
        T1B=DLOG(F2)
        T1BSQ=(DLOG(F2))**2.
        T2B=F2**(I-1)
        T3B=(1.-F2)**(M-I)
        B=T1B*T2B*T3B
        BSQ=T1BSQ*T2B*T3B
        EL(I)=EL(I)+(A+B)*C/2.
        ELSQ(I)=ELSQ(I)+(ASQ+BSQ)*C/2.
        GO TO 10
15      EL(I)=EL(I)*D
        ELSQ(I)=ELSQ(I)*D
        W(I)=ELSQ(I)-(EL(I))**2.
        W(I)=1./W(I)
        WRITE(6,100)EL(I),ELSQ(I),W(I)
100     FORMAT(1X,3F30.12)
        D=D*(M-I)/I
1       CONTINUE
        WRITE(2,200)(EL(I),I=1,M)
        WRITE(3,200)(W(I),I=1,M)
200     FORMAT(1X,F40.15)
C       THE EL(I) AND W(I) ARE STORED ON DISK FOR FUTURE USE.
        LOCK 2
        LOCK 3
        STOP
        END
```

Appendix C

Program FLOOD

```
C      THIS PROGRAM FINDS ESTIMATES FOR THE PARAMETERS ALPHA(1),
C      ALPHA(2), THETA(1), AND THETA(2)
C                      REQUIRED INPUT INCLUDES M, THE NUMBER OF
C      YEARS OF THE DATA RECORD, BB, THE MAXIMUM POSSIBLE FLOOD
C      HEIGHT, AND CC, THE SCALE FACTOR.  THE ORDERED FLOOD DATA,
C      THE EL(I), AND THE Z(I) ARE READ INTO ARRAYS X, EL, AND W
C      RESPECTIVELY (THE W(I) ARE COMPUTED AND STORED IN ARRAY W
C      DURING EXECUTION).  THE MINIMIZATION PROCESS IS ACHIEVED WITH
C      A SUBROUTINE FROM THE IMSL (1977) LIBRARY CALLED ZSYSTM

C                    THIS SUBROUTINE REQUIRES AN EXTERNAL FUNCTION (F),
C      TWO CONVERGENCE CRITERIA (EPS AND NSIG), THE NUMBER OF
C      UNKNOWNS (N), THE MAXIMUM NUMBER OF ITERATIONS OF THE
C      EXTERNAL FUNCTION F (ITMAX), A WORK AREA OF COMPUTER
C      STORAGE (WA), AN ARRAY FOR PASSING PARAMETERS (PAR, WHICH
C      IS NOT USED IN THIS STUDY), AN ERROR MESSAGE VARIABLE (IER),
C      AND STARTING VALUES FOR THE ALPHAS.  THE STARTING VALUES
C      FOR ALPHA(1) AND ALPHA(2) ARE COMPUTED FROM THE ORDERED
C      DATA.  OUTPUT CONSISTS OF ALHPA(1), ALPHA(2), THETA(1),
C      THETA(2), ITMAX, AND IER, THE ERROR MESSAGE.  IER=0 MEANS
C      THERE ARE NO ERRORS AND MINIMIZATION WAS COMPLETED TO THE
C      ACCURACY SPECIFIED BY THE CONVERGENCE CRITERIA.
C      FOR MORE DETAILED INFORMATION ON THE SUBROUTINE ZSYSTEM,
C      SEE THE IMSL (1977) LIBRARY.
FILE   1(KIND=DISK,TITLE="(878073)SAGUENAY/DATA")
FILE   2(KIND=DISK,TITLE="(878073)SAGUENAY/EL")
FILE   3(KIND=DISK,TITLE="(878073)SAGUENAY/W")

       EXTERNAL F
       DIMENSION ALPHA(2),WA(20),PAR(2),XREG(200),YREG(200)
       COMMON M,BB,CC,X(200),W(200),EL(200),THETA(2),Y(200)
       EPS=1.0E-9
       NSIG=5
       N=2
       ITMAX=100
       IER=0
C      M--THE NUMBER OF YEARS OF DATA, BB--THE MAXIMUM POSSIBLE FLOOD
C      HEIGHT, AND CC--THE SCALE FACTOR ARE READ.
       READ(5,/)M
       READ(5,/)BB,CC
C      THE ORDERED DATA, THE EL(I), AND THE Z(I) ARE READ INTO ARRAYS
C      X, EL, AND W RESPECTIVELY.
       READ(1,101)(X(I),I=1,M)
101    FORMAT(1X,F12.2)
       READ(2,200)(EL(I),I=1,M)
       READ(3,200)(W(I),I=1,M)
200    FORMAT(1X,F40.15)
C      THE W(I) ARE CALCULATED.
       DO 23 I=1,M
       W(I)=W(I)/W(M)
23     CONTINUE
C      STARTING VALUES ARE DETERMINED FOR ALPHA(1) AND ALPHA(2).
       SUMX1=0.
       SUMY1=0.
       SUMXY1=0.
       SUMXX1=0.
       DO 15 I=1,4
       XREG(I)=ALOG(BB-X(I))
       YREG(I)=ALOG(-ALOG(I/(M+1.)))
```

```
      SUMX1=SUMX1+XREG(I)
      SUMY1=SUMY1+YREG(I)
      SUMXY1=SUMXY1+XREG(I)*YREG(I)
      SUMXX1=SUMXX1+XREG(I)**2
15    CONTINUE
      X1BAR=SUMX1/4.
      Y1BAR=SUMY1/4.
      ALPHA(1)=(SUMXY1-4.*X1BAR*Y1BAR)/(SUMXX1-4.*X1BAR**2)
      SUMX2=0.
      SUMY2=0.
      SUMXY2=0.
      SUMXX2=0.
      DO 16 J=M-3,M
      XREG(J)=ALOG(BB-X(J))
      YREG(J)=ALOG(-ALOG(J/(M+1.)))
      SUMX2=SUMX2+XREG(J)
      SUMY2=SUMY2+YREG(J)
      SUMXY2=SUMXY2+XREG(J)*YREG(J)
      SUMXX2=SUMXX2+XREG(J)**2
16    CONTINUE
      X2BAR=SUMX2/4.
      Y2BAR=SUMY2/4.
      ALPHA(2)=(SUMXY2-4.*X2BAR*Y2BAR)/(SUMXX2-4.*X2BAR**2)
      WRITE(6,50)M,BB,CC
50    FORMAT(1X,'THE NUMBER OF YEARS OF THE DATA RECORD=',I15,//,
     *1X,'THE MAXIMUM POSSIBLE FLOOD HEIGHT=',I15,//,1X,
     * 'THE SCALE FACTOR=',I15,//)
      WRITE(6,60)ALPHA(1),ALPHA(2)
60    FORMAT(1X,'THE STARTING VALUES ARE',/,1X,'ALPHA(1)=',F15.5,
     * 5X,'ALPHA(2)=',F15.5,/////)
C     ZSYSTM IS CALLED TO MINIMIZE EQUATION (20) AND OUTPUT THE
C     ESTIMATED PARAMETERS.
      CALL ZSYSTM(F,EPS,NSIG,N,ALPHA,ITMAX,WA,PAR,IER)
      WRITE(6,70)ITMAX,IER
70    FORMAT(1X,'NUMBER OF ITERATIONS OF EXTERNAL FUNCTION=',I5,
     * //,1X,'ERROR MESSAGE=',I5,/////)
      WRITE(6,80)
80    FORMAT(1X,'PARAMETER ESTIMATES ARE',//)
      WRITE(6,90)ALPHA(1),ALPHA(2),THETA(1),THETA(2)
90    FORMAT(1X,'ALPHA(1)=',F20.10,5X,'ALPHA(2)=',F20.10,//,1X,
     * 'THETA(1)=',F20.10,5X,'THETA(2)=',F20.10)
      STOP
      END
```

##

```
      FUNCTION F(ALPHA,KK,PAR)
C     THIS FUNCTION EVALUATES THE TWO EQUATIONS IN TWO UNKNOWNS.
      DIMENSION WEL(200),WLX(200),WELX(200),ALPHA(2),PAR(2)
      COMMON M,BB,CC,X(200),W(200),EL(200),THETA(2),Y(200)
      DO 10 I=1,M
      WEL(I)=W(I)*EL(I)
      Y(I)=(BB-X(I))/CC
      WLX(I)=W(I)*ALOG(Y(I))
10    WELX(I)=WLX(I)*EL(I)
      A3=2.*ALPHA(1);A4=2.*ALPHA(2);A5=ALPHA(1)+ALPHA(2)
      Z1=0.;Z2=0.;Y1=0.;Y2=0.;Y3=0.;Y4=0.;Y5=0.;Y6=0.;Y7=0.;Y8=0.
      B1=0.;B2=0.;B3=0.;B4=0.;B5=0.
      DO 20 I=1,M
      YA1=Y(I)**ALPHA(1)
      YA2=Y(I)**ALPHA(2)
```

```
      YA3=Y(I)**A3
      YA4=Y(I)**A4
      YA5=Y(I)**A5
      Z1=Z1-WEL(I)*YA1
      Z2=Z2-WEL(I)*YA2
      Y1=Y1+W(I)*YA3
      Y2=Y2+W(I)*YA4
      Y3=Y3+W(I)*YA5
      Y4=Y4+WELX(I)*YA1
      Y5=Y5+WLX(I)*YA3
      Y6=Y6+WLX(I)*YA5
      Y7=Y7+WELX(I)*YA2
20    Y8=Y8+WLX(I)*YA4
      TH1=(Z1*Y2-Z2*Y3)/(Y1*Y2-Y3*Y3)
      TH1=ABS(TH1)
      TH2=(Z1-TH1*Y1)/Y3
      TH2=ABS(TH2)
      ALPHA(1)=ABS(ALPHA(1))
      ALPHA(2)=ABS(ALPHA(2))
      GO TO (55,56),KK
55    F=Y4+TH1*Y5+TH2*Y6
      THETA(1)=TH1
      THETA(2)=TH2
      RETURN
56    F=Y7+TH1*Y6+TH2*Y8
      THETA(1)=TH1
      THETA(2)=TH2
      RETURN
      END
```

##

Appendix D

Program PFHTS

```
C-   THIS PROGRAM IS AN INTERACTIVE (TERMINAL) PROGRAM.
C-   GIVEN A DISTRIBUTION FUNCTION F(X) OF THE FORM OF
C-   EQUATION (8) OF CHAPTER 2 WHERE C AND C' HAVE BEEN
C-   REPARAMETERIZED AS THETA(1) AND THETA(2) AS IN
C-   EQUATION (20) OF CHAPTER 3, FOR ANY X (FLOOD HEIGHT)<
C-   F(X) THE PROBABILITY THAT ANY POSSIBLE FLOOD IS
C-   LESS THAN OR EQUAL TO X IS EVALUATED.  THUS THE
C-   SELECTED RETURN PERIODS OR RECURRANCE INTERVALS
C-   CAN BE FOUND BY FINDING SOME X (TO THE NEAREST
C-   INTEGER) WHICH YEILDS THE DESIRED F(X) PROBABILITY.
C-   REQUIRED AS INPUT ARE BB--THE MAXIMUM POSSIBLE
C-   FLOOD HEIGHT, CC--THE SCALE FACTOR, ALPHA(1), ALPHA(2),
C-   THETA(1), AND THETA(2)--THE PARAMETERS OF THE
C-   PARTICULAR DISTRIBUTION FUNCTION F(X), AND X--THE
C-   FLOOD HEIGHT FOR WHICH F(X) IS DESIRED.  F(X) MAY
C-   BE FOUND FOR AS MANY X VALUES AS REQUIRED.
C-   WHEN FINISHED SIMPLY ENTER   ?END  AND A NEW
C-   F(X) MAY BE EXAMINED OR ONE MAY LOG OFF THE
C-   COMPUTER AS DESIRED.
DIMENSION ALPHA(2),THETA(2)
 WRITE(6,160)
 160 FORMAT( 1X," ENTER BB AND CC")
READ(5,/)BB,CC
 WRITE(6,170)
 170 FORMAT( 1X,"ENTER ALPHA(1) AND ALPHA(29",/)
READ(5,/)(ALPHA(I),I=1,2)
 WRITE(6,180)
 180 FORMAT( 1X," ENTER THETA(1) AND THETA(2)",?)
READ(5,/)(THETA(I),I=1,2)
 1 WRITE(6,190,END=99)
 190 FORMAT( 1X," ENTER X",/)
READ(5,/,END=99)X
Y=(BB-X)/CC
 F=EXP(-THETA(1)*Y**ALPHA(1)-THETA(2)*Y:*ALPHA(2))
 WRITE(6,100)F
 100 FORMAT(1X,'F(X)=',F20.15,/)
GO TO 1
99 STOP
END
```

Appendix E

Program EPROB

```
C      THIS PROGRAM CALCULATES THE EXPECTED PROBABILITY OF A
C      MAXIMUM YEARLY FLOOD BEING GREATER THAN OR EQUAL TO A GIVEN
C      HEIGHT USING THE OBSERVED DATA RECORD.  THESE PROBABILITIES
C      ARE ESTIMATED USING THE THREE FORMULAE
C      -- HAZEN, CHEGODAYEV, AND WEIBULL.  THE DESIRED RECURRANCE
C      INTERVALS OR RETURN PERIODS ARE FOUND BY LINEAR INTER-
C      POLATION BETWEEN THE TWO OBSERVED FLOOD HEIGHTS WHOSE
C      EXPECTED PROBABILITIES BRACKET THE DESIRED PROBABILITY.
C      REQUIRED INPUT IS THE NUMBER OF YEARS OF THE DATA RECORD
C      AND THE DATA ITSELF.  M IS THE NUMBER OF YEARS OF DATA AND
C      X IS AN ARRAY FOR THE FLOOD RECORD.
FILE   1(KIND=DISK,TITLE="SAGUENAY/DATA")

       DIMENSION X(200),HAZPR(200),CHEGPR(200),WEIBPR(200)
C      M, THE NUMBER OF YEARS OF DATA, IS READ.
       READ(5,/)M

C      THE FLOOD DATA IS READ INTO ARRAY X FROM DISK.
       READ(1,101)(X(I),I=1,M)
101    FORMAT(1X,F12.2)
C      EXPECTED PROBABILITIES ARE CALCULATED.  HAZPR, CHEGPR,
C      AND WEIBPR ARE ARRAYS FOR THE PROBABILITIES FOUND USING THE
C      HAZEN, CHEGODAYEV, AND WEIBULL FORMULAE RESPECTIVELY.
       ITEST=M/2.-1.
       DO 105 I=ITEST,M
       HAZPR(I)=((M-I+1.)-0.5)/M
       CHEGPR(I)=((M-I+1.)-0.3)/(M+0.4)
       WEIBPR(I)=(M-I+1.)/(M+1.)
105    CONTINUE
       WRITE(6,200)

200    FORMAT(1X,' EXPECTED PROBABILITIES',/////)
       WRITE(6,99)
99     FORMAT(1X,5X,'DATA',32X,' HAZEN',25X,'CHEGODAYEV',19X,' WEIBULL')
       WRITE(6,100) (X(I),HAZPR(I),CHEGPR(I),WEIBPR(I),I=ITEST,M)
100    FORMAT(1X,F12.2,18X,3F30.16)
       STOP
       END
```

Appendix F

Program DEVIATION

```
C      THIS PROGRAM COMPUTES THE MEAN OF THE ABSOLUTE RELATIVE DEVI-
C      ATIONS AND THE MEAN QUADRATIC DEVIATION BETWEEN A GIVEN
C      DATA SET AND ITS PREDICTING DISTRIBUTION FUNCTION FOR A
C      SELECTED SET OF RETURN PERIODS AS DESCRIBED IN CHAPTER 4.
C      HAZEN, CHEGODAYEV, AND WEIBULL ARE TREATED AS DIFFERENT
C      METHODS. REQUIRED INPUT INCLUDES TH, TC, AND TW, THE NUMBER
C      OF SELECTED RECURRANCE INTERVALS FOR THE HAZEN, CHEGODAYEV,
C      AND WEIBULL METHODS RESPECTIVELY. THE PREDICTED FLOOD HEIGHTS
C      OF THE ESTIMATED DISTRIBUTION FUNCTION ARE OBTAINED FROM
C      THE INTERACTIVE PROGRAM PFHTS AND ARE INPUT AS ARRAY PF.
C      THE EXPECTED FLOOD HEIGHTS FOR THE SELECTED RETURN PERIODS
C      ARE FOUND USING THE PROGRAM PROB AND LINEAR INTERPOLATION
C      AND ARE INPUT AS ARRAYS EFH, EFC, AND EFW FOR THE HAZEN,
C      CHEGODAYEV AND WEIBULL METHODS RESPECTIVELY.

       DIMENSION EFH(10),EFC(10),EFW(10),PF(10),H(10),C(10),W(10),PR(10)
       PR(1)=.5;PR(2)=.8;PR(3)=.9;PR(4)=.95;PR(5)=.98;PR(6)=.99
C      TH, TC, AND TW, THE NUMBER OF RETURN PERIODS, ARE READ.
       READ(5,/)TH,TC,TW

C      THE PREDICTED FLOOD HEIGHTS ARE READ INTO ARRAY PF.
       READ(5,/) (PF(I),I=1,6)
C      THE EXPECTED FLOOD HEIGHTS ARE READ INTO ARRAYS EFH, EFC,
C      AND EFW RESPECTIVELY.
       READ(5,/) (EFH(I),I=1,TH)
       READ(5,/) (EFC(I),I=1,TC)
       READ(5,/) (EFW(I),I=1,TW)
       WRITE(6,10)

10     FORMAT( 1X,' PROBABILITY',2X,' PREDICTED HEIGHT',5X,'HAZEN',8X,
     * ' CHEGODAYEV',6X,' WEIBULL',/)
       WRITE(6,20)(PR(I),PF(I),EFH(I),EFC(I),EFW(I),I=1,6)
20     FORMAT( 1X,F10.2,8X,F8.2,8X,F8.2,8X,F8.2,8X,F8.2)
C      MEAN ABSOLUTE AND MEAN QUADRATIC DEVIATIONS ARE COMPUTED
C      FOR EACH METHOD EMPLOYING THREE DO LOOPS USING EQUATIONS (23)
C      AND (24) OF CHAPTER 4.  THE SMALLER THE DEVIATIONS THE BETTER
C      THE FIT.
       DIFFH=0.;DIFFC=0.;DIFFW=0.;DHS=0.;DCS=0.;DWS=0.
       DO 1 I=1,TH
       H(I)=ABS((PF(I)-EFH(I))/EFH(I)*100.)
       DIFFH=DIFFH+H(I)
       DHS=DHS+H(I)*H(I)
1      CONTINUE
       DO 2 I=1,TC
       C(I)=ABS((PF(I)-EFC(I))/EFC(I)*100.)
       DIFFC=DIFFC+C(I)
       DCS=DCS+C(I)*C(I)
2      CONTINUE
       DO 3 I=1,TW
       W(I)=ABS((PF(I)-EFW(I))/EFW(I)*100.)
       DIFFW=DIFFW+W(I)
       DWS=DWS+W(I)*W(I)
3      CONTINUE
       ADIFFH=DIFFH/TH
       ADIFFC=DIFFC/TC
       ADIFFW=DIFFW/TW
       ADHS=DHS/TH
       ADCS=DCS/TC
       ADWS=DWS/TW
       WRITE(6,100) ADIFFH,ADIFFC,ADIFFW
100    FORMAT(////,' MEAN OF THE ABSOLUTE RELATIVE DEVIATIONS',///,10X,
     * ' HAZEN',F20.2 ,/, 5X,' CHEGODAYEV',F20.2 ,/, 8X,' WEIBULL',
     * F20.2 ,////)
       WRITE(6,200) ADHS,ADCS,ADWS
200    FORMAT( 5X ,' MEAN QUADRATIC DEVIATION',///, 10X,' HAZEN',
     * F20.2 ,/, 6X, 'CHEGODAYEV',F20.2 ,/, 9X,'WEIBULL',F20.2 )
       STOP
       END
```

Appendix G

Program MAXIMUM LIKELIHOOD

```
      DOUBLE PRECISION T(100),THETA(550),EK(550),X(56),Y(55),SL,SLK,ELNM
     1,R,ANG,EI,ZRK,C(550),SK
C     INPUT
C     N=SAMPLE SIZE (BEFORE CENSORING),N=100 OR LESS AS DIMENSIONED
C     SS1=0 IF SCALE PARAMETER THETA IS KNOWN
C     SS1=1 IF SCALE PARAMETER THETA IS TO BE ESTIMATED
C     SS2=0 IF SHAPE PARAMETER K IS KNOWN
C     SS2=1 IF SHAPE PARAMETER K IS TO BE ESTIMATED
C     SS3=0 IF LOCATION PARAMETER C IS KNOWN
C     SS3=1 IF LOCATION PARAMETER C IS TO BE ESTIMATED
C     T(I)=I-TH ORDER STATISTIC OF SAMPLE (I=1,N)
C      (SUBSTITUTE BLANK CARDS FOR UNKNOWN CENSORED OBSERVATIONS)
C     M=NUMBER OF OBSERVATIONS REMAINING AFTER CENSORING N-M FROM ABOVE
C     C(1)=INITIAL ESTIMATE (OR KNOWN VALUE) OF C
C     THETA(1)=INITIAL ESTIMATE (OR KNOWN VALUE) OF THETA
C     EK(1)=INITIAL ESTIMATE (OR KNOWN VALUE) OF K
C     MR=NUMBER OF OBSERVATIONS CENSORED FROM BELOW, NORMALLY 0 INITIALL
C     OUTPUT
C     N,SS1,SS2,SS3,M,C(1),THETA(1),EK(1),MR--SAME AS FOR INPUT
C     C(J)=ESTIMATE AFTER J-1 ITERATIONS (OR KNOWN VALUE) OF C
C     THETA(J)=ESTIMATE AFTER J-1 ITERATIONS (OR KNOWN VALUE) OF THETA
C     EK(J)= ESTIMATE AFTER J-1 ITERATIONS (OR KNOWN VALUE) OF K
C      (MAXIMUM VALUE OF J AS PRESENTLY DIMENSIONED IS 550)
C     EL=NATURAL LOGARITHM OF LIKELIHOOD FOR C(J), THETA(J),EK(J)
C     REFERENCE
C     HARTER,H. LEON AND MOORE, ALBERT H., MAXIMUM-LIKELIHOOD ESTI-
C      MATION OF THE PARAMETERS OF GAMMA AND WEIBULL POPULATIONS
C      FROM COMPLETE AND FROM CENSORED SAMPLES, TECHNOMETRICS,
C      7 (1965), 639-643.  ERRATA,9 (1967), 195
      IF(N) 66,66,77
   77 EN=N
      IF(M) 64,64,32
   32 EM=M
 31   ELNM=0.
      EMR=MR
      MRP=MR+1
 33   NM=N-M+1
      DO 34 I=NM,N
      EI=I
 34   ELNM=ELNM+DLOG(EI)
      IF(MR) 66,35,74
 74   DO 75 I=1,MR
      EI=I
 75   ELNM=ELNM-DLOG(EI)
 35   DO 30 J=1,550
      IF(J-1) 66,25,37
 37   JJ=J-1
      SK=0.
      SL=0.
      DO 6 I=MRP,M
 6    SK=SK+(T(I)-C(JJ))**EK(JJ)
      IF(SS1)7,7,8
 7    THETA(J)=THETA(JJ)
      GO TO 9
 8    IF(MR) 66,19,20
 19   THETA(J)=((SK+(EN-EM)*(T(M)-C(JJ))**EK(JJ))/EM)**(1./EK(JJ))
      GO TO 9
 20   X(1)=THETA(JJ)
      LS=0
```

```
      DO 21 L=1,55
      LL=L-1
      LP=L+1
      X(LP)=X(L)
      ZRK=((T(MRP)-C(JJ))/X(L))**EK(JJ)
      Y(L)=-EK(JJ)+(EM-EMR)/X(L)+EK(JJ)*SK/X(L)**(EK(JJ)+1.)+EK(JJ)*
     1(EN-EM)*(T(M)-C(JJ))**EK(JJ)/X(L)**(EK(JJ)+1.)-EMR*EK(JJ)*ZRK*
     2DEXP(-ZRK)/(X(L)*(1.-DEXP(-ZRK)))
      IF(Y(L)) 53,73,54
53    LS=LS-1
      IF (LS+L) 58,55,58
54    LS=LS+1
      IF (LS-L) 58,56,58
55    X(LP)=.5*X(L)
      GO TO 61
56    X(LP)=1.5*X(L)
      GO TO 61
58    IF(Y(L)*Y(LL)) 60,73,59
59    LL=LL-1
      GO TO 58
60    X(LP)=X(L)+Y(L)*(X(L)-X(LL))/(Y(LL)-Y(L))
 61   IF(DABS(X(LP)-X(L))-1.E-4) 73,73,21
21    CONTINUE
73    THETA(J)=X(LP)
9     EK(J)=EK(JJ)
10    IF(SS2) 12,12,11
11    DO 17 I=MRP,M
17    SL=SL+DLOG(T(I)-C(JJ))
      X(1)=EK(J)
      LS=0
      DO 51 L=1,55
      SLK=0.
      DO 18 I=MRP,M
18    SLK=SLK+(DLOG(T(I)-C(JJ))-DLOG(THETA(J)))*(T(I)-C(JJ))**X(L)
      LL=L-1
      LP=L+1
      X(LP)=X(L)
      ZRK=((T(MRP)-C(JJ))/THETA(J))**X(L)
      Y(L)=(EM-EMR)*(1./X(L)-DLOG(THETA(J)))+SL-SLK/THETA(J)**
     -X(L)+(EN-EM)*(DLOG(THETA(J))-
     +DLOG(T(M)-C(JJ)))*(T(M)-C(JJ))**X(L)/
     2THETA(J)**X(L)+EMR*ZRK*(DLOG(ZRK)/X(L))*DEXP(-ZRK)/
     3(1.-DEXP(-ZRK))
      IF(Y(L)) 43,52,44
43    LS=LS-1
      IF(LS+L) 47,45,47
44    LS=LS+1
      IF(LS-L) 47,46,47
45    X(LP)=.5*X(L)
      GO TO 50
46    X(LP)=1.5*X(L)
      GO TO 50
47    IF(Y(L)*Y(LL)) 49,52,48
48    LL=LL-1
      GO TO 47
49    X(LP)=X(L)+Y(L)*(X(L)-X(LL))/(Y(LL)-Y(L))
 50   IF (DABS(X(LP)-X(L))-1.E-4) 52,52,51
51    CONTINUE
52    EK(J)=X(LP)
12    C(J)=C(JJ)
62    IF(SS3) 25,25,14

14    IF(1.-EK(J)) 16,78,78
78    IF(SS1+SS2) 57,57,16
16    X(1)=C(J)
      LS=0
      DO 23 L=1,55
```

```
      SK1=0.
      SR=0.
      DO 15 I=MRP,M
      SK1=SK1+(T(I)-X(L))**(EK(J)-1.)
15    SR=SR+1.*(T(I)-X(L))
      LL=L-1
      LP=L+1
      X(LP)=X(L)
      ZRK=((T(MRP)-X(L))/THETA(J))**EK(J)
      Y(L)=(1.-EK(J))*SR+EK(J)*(SK1+(EN-EM)*(T(M)-X(L))**(EK(J)-1.))
     1/THETA(J)**EK(J)-EMR*EK(J)*ZRK*DEXP(-ZRK)/((T(MRP)-X(L))*(1.-
     1DEXP(-ZRK)))
      IF(Y(L)) 39,24,40
39    LS=LS-1
      IF(LS+L) 70,41,70
40    LS=LS+1
      IF(LS-L) 70,42,70
41    X(LP)=.5*X(L)
      GO TO 22
42    X(LP)=.5*X(L)+.5*T(1)
      GO TO 22
70    IF(Y(L)*Y(LL)) 72,24,71
71    LL=LL-1
      GO TO 70
72    X(LP)=X(L)+Y(L)*(X(L)-X(LL))/(Y(LL)-Y(L))
   22 IF(DABS(X(LP)-X(L))-1.E-4)24,24,23
23    CONTINUE
24    C(J)=X(LP)
      GO TO 25
57    C(J)=T(1)
25    IF(MR) 66,38,69
38    DO 63 I=1,M
      IF(C(J)+1.E-4-T(I)) 68,67,67
67    MR=MR+1
63    C(1)=T(1)
68    IF(MR) 66,69,31
69    SK=0.
      SL=0.
      DO 36 I=MRP,M
      SK=SK+(T(I)-C(J))**EK(J)
36    SL=SL+DLOG(T(I)-C(J))
      ZRK=((T(MRP)-C(J))/THETA(J))**EK(J)
      FL=FLNM+(EM-EMR)*(DLOG(EK(J))-EK(J)*DLOG(THETA(J)))+(EK(J)-1.)*SL-
     1(SK+(EN-EM)*(T(M)-C(J))**EK(J))/(THETA(J)**EK(J))+EMR*DLOG(1.-DEXP
     2(-ZRK))
      IF(J-3) 30,27,27
   27 IF(DABS(C(J)-C(JJ))-1.E-4) 28,28,30
   28 IF  (DABS(THETA(J)-THETA(JJ))-1.E-4) 29,29,30
   29 IF(DABS(EK(J)-EK(JJ))-1.E-4)100,100,30
30    CONTINUE
   64 PRINT /,'      NO YIELD POINT'
      RETURN
  100 R=EK(J)
      ANG=THETA(J)
      RETURN
66    STOP




      END
```

Appendix H

Data Used in Analysis

STATION	COUNTRY	RIVER	LOCATION
hB24	SENEGAL	SENEGAL	BAKEL

1040	1740	1880	2290	2750	2850	2850	2890
3140	3290	3320	3400	3480	3550	3560	3560
3500	3600	3600	3760	3770	3840	3840	4180
4200	4200	4300	4350	4400	4460	4620	4620
4680	4790	4850	4970	5070	5260	5330	5330
5430	5450	5450	5450	5450	5450	5590	5590
5620	6030	6310	6410	6430	6570	6640	7000
7030	7180	7300	7600	7630	8170	9070	9940

STATION	COUNTRY	RIVER	LOCATION
hE60	U.S.A	SUSQUEHANNA	HARRIABURA, PA.

3850	4330	4390	5010	5012	5040	5100	5150
[illegible]	6000	6060	6116	6230	6460	6500	6513
6540	6650	6850	6853	6910	6940	6990	7050
7050	7051	7079	7140	7150	7390	7500	7500
7620	7646	7650	7820	7870	7957	8100	8160
8210	8330	8410	8410	8440	8670	8720	8920
9160	9170	9170	9175	9400	9571	10100	10700
10730	10817	11100	11400	11600	11700	11780	11800
12000	12700	13705	14000	17400	21000		

STATION	COUNTRY	RIVER	LOCATION
IB06	INDIA	KRISHNA	VIJAYAWADA

7190	9058	9915	10017	10204	10212	10360	10458
10478	10495	10613	10793	10813	10878	10882	10916
11105	11122	11374	11500	12091	12399	12560	12912
12979	13069	13113	13260	13465	13528	13582	13686
14033	14132	14220	14242	14503	14520	15396	15514
15647	15816	15872	16009	16380	16524	16782	17372
17680	17908	17970	18511	18888	19879	20970	23501
25902	26873	27073	29768				

STATION	COUNTRY	RIVER	LOCATION
BF40	CZECHOSLOVAKIA	ELBE	DECIN

543	587	595	610	726	1038	1046	1058
1112	1117	1138	1138	1149	1160	1166	1172
1175	1181	1181	1198	1205	1207	1234	1246
1265	1265	1269	1270	1282	1293	1300	1312
1317	1350	1354	1360	1372	1396	1429	1454
1462	1474	1492	1498	1522	1527	1546	1561
1565	1565	1575	1601	1610	1618	1643	1702
1717	1742	1768	1845	1848	1853	1874	1915
1930	1930	1940	2038	2040	2040	2083	2109
2124	2146	2158	2250	2284	2301	2373	2379
2385	2400	2410	2515	2540	2565	2600	2626
2643	2666	2725	2815	2850	2876	2937	2937
2940	2975	3100	3172	3343	3600	3770	3779
4058	4143	4450	4822				

STATION	COUNTRY	RIVER	LOCATION
BE38	GERMANY	DANUBE	HOFKIRCHEN

947	956	1090	1090	1100	1120	1230	1230
1250	1250	1260	1260	1310	1310	1320	1320
1340	1350	1380	1400	1440	1450	1450	1460
1460	1480	1540	1560	1600	1640	1650	1720
1730	1760	1800	1810	1810	1850	1850	1880
1890	1900	1920	1930	1980	2020	2030	2040
2050	2070	2150	2170	2180	2240	2270	2310
2390	2400	2450	2540	2600	2690	2780	2780
2810	2930	3000	3880				

STATION	COUNTRY	RIVER	LOCATION
BF19	NORWAY	GLOMA	LANGNES

1157	1267	1351	1358	1413	1504	1504	1518
1533	1557	1566	1580	1643	1650	1675	1707
1734	1738	1770	1783	1817	1822	1839	1872
1878	1910	1916	1953	2031	2050	2050	2100
2106	2133	2168	2172	2180	2195	2232	2240
2255	2256	2258	2260	2288	2299	2302	2311
2312	2321	2346	2359	2363	2380	2385	2390
2515	2582	2585	2715	2850	2877	3160	3224
3429	3543						

STATION	COUNTRY	RIVER	LOCATION
CF25	USSR	NEMAN	SMALININKAI

810	870	980	1050	1100	1150	1150	1200
1240	1250	1300	1350	1400	1400	1400	1400
1450	1500	1550	1550	1600	1600	1600	1650
1650	1700	1700	1700	1700	1700	1750	1750
1750	1800	1800	1800	1800	1850	1850	1900
1900	1950	1950	1950	1950	1950	2000	2000
2000	2000	2100	2100	2100	2100	2100	2100
2100	2100	2100	2100	2100	2200	2200	2200
2300	2300	2300	2300	2300	2300	2300	2300
2400	2400	2400	2400	2400	2500	2500	2500
2500	2500	2600	2600	2600	2600	2600	2600
2700	2700	2700	2700	2700	2700	2700	2700
2700	2800	2800	2800	2800	2900	2900	2900
3000	3000	3000	3000	3000	3000	3000	3000
3100	3100	3100	3100	3200	3200	3200	3200
3200	3200	3300	3400	3400	3400	3400	3400
3500	3500	3600	3600	3600	3700	3700	3800
3900	3900	4100	4200	4300	4300	4300	4600
4600	4700	4800	4900	5200	5600	5800	6200
6200	6600	6800					

STATION	COUNTRY	RIVER	LOCATION
mE19	CANADA	FRASER	HOPE

5130	5810	6000	6060	6830	7080	7220	7220
7420	7480	7560	7620	7700	7820	7820	7820
7840	7900	8040	8040	8040	8160	8210	8330
8470	8500	8500	8520	8550	8580	8670	8670
8720	8840	8980	9010	9060	9260	9290	9350
9520	9540	9690	9690	9770	9770	9910	9970
10300	10300	10500	10600	10800	10800	11100	11300
11600	12500	15200					

STATION		COUNTRY		RIVER		LOCATION	
JE792		CANADA		ASSINIBOINE		HEADINGLEY	
49	54	61	62	65	92	114	116
117	129	139	146	146	153	174	185
191	202	204	206	216	216	217	222
228	230	233	236	248	264	269	275
275	281	286	289	292	300	306	317
320	340	346	360	382	388	430	473
473	481	519	547	564	566	592	595
615							

STATION		COUNTRY		RIVER		LOCATION	
1F00		CANADA		S. SASKATCHEWAN		MEDICINE HAT	
230	317	379	391	524	572	575	581
649	683	683	688	722	725	731	733
821	824	827	899	912	940	940	952
957	960	963	974	983	991	991	1030
1040	1040	1070	1090	1090	1090	1130	1290
1370	1520	1550	1630	1690	1830	1840	1880
2080	2090	2170	2200	2400	2550	2710	3060
3710	4080						

STATION		COUNTRY		RIVER		LOCATION	
KF62		CANADA		S. SASKATCHEWAN		SASKATOON	
399	541	583	583	595	632	793	816
852	855	855	861	901	926	980	994
1050	1070	1070	1080	1110	1120	1140	1150
1170	1180	1190	1210	1250	1260	1270	1280
1370	1370	1420	1420	1420	1420	1530	1540
1540	1570	1630	1760	1780	1820	1850	1970
2180	2330	2420	2490	2630	2700	3060	3140
3140	3370	3940					

STATION		COUNTRY		RIVER		LOCATION	
KF53		CANADA		N. SASKATCHEWAN		PRINCE ALBERT	
487	527	589	620	623	683	685	756
759	762	765	770	790	796	799	875
926	940	952	954	991	1010	1010	1050
1070	1110	1120	1130	1140	1180	1190	1200
1230	1250	1250	1270	1280	1340	1350	1510
1540	1560	1570	1570	1570	1620	1620	1640
1650	1790	1800	1980	2090	2160	2460	2790
2930	2970	5300					

STATION		COUNTRY		RIVER		LOCATION	
HE88a		CANADA		HURRICANA		AMOS	
99	99	117	118	125	132	132	135
142	146	150	154	158	158	161	161
161	164	164	166	167	172	172	173
173	174	179	183	183	185	192	194
195	195	201	202	204	205	213	213
216	229	230	230	235	240	244	262
262	264	283	317	337			

STATION		COUNTRY		RIVER		LOCATION	
JF50a		CANADA		WINNIPEG		SLAVE FALLS	
666	668	668	901	986	1000	1020	1030
1050	1060	1060	1090	1100	1140	1200	1250
1250	1270	1290	1370	1390	1420	1450	1460
1510	1590	1720	1720	1750	1790	1920	1970
1990	2040	2190	2260	2390	2410	2450	2780
2800							

STATION		COUNTRY		RIVER		LOCATION	
HE1833		CANADA		SAGUENAY		ISLE-MALIGNE	
2370	2380	2410	2730	2830	3400	3510	3600
3650	3770	3820	3850	3850	3860	4050	4050
4080	4110	4190	4190	4250	4420	4420	4420
4450	4530	4530	4590	4640	4670	4670	4870
4930	4930	4950	5010	5070	5150	5180	5270
5550	5660	5720	5830	5920	6030	6120	6370
6460	6460	6480	6740	6770	6820	7390	7930
9060	9260						

STATION		COUNTRY		RIVER		LOCATION	
AB36		MALI		NIGER		DIRE	
1947	1965	2001	2061	2061	2120	2139	2145
2157	2199	2205	2217	2223	2223	2262	2269
2279	2300	2308	2314	2314	2321	2335	2359
2384	2384	2392	2405	2405	2411	2418	2431
2440	2447	2535	2557	2565	2595	2625	2632
2640	2647	2655	2677	2677			

STATION		COUNTRY		RIVER		LOCATION	
AB72		MALI		NIGER		KOULIKORO	
3646	4010	4290	4467	4830	4920	4920	4980
4980	5000	5140	5186	5240	5285	5375	5375
5437	5505	5580	5610	5624	5670	5760	5790
5910	6002	6170	6172	6210	6220	6220	6280
6360	6380	6420	6440	6440	6480	6540	6550
6640	6740	6840	6900	6940	6946	6960	6960
6980	7020	7228	7247	7400	7456	7560	7610
7740	7798	8740	9300	9500	9700		

STATION		COUNTRY		RIVER		LOCATION	
AE85		USA		PENOBSCOT		WEST ENFIELD,ME.	
821	903	917	928	957	1000	1040	1120
1130	1150	1175	1190	1220	1250	1270	1331
1360	1380	1420	1436	1440	1460	1520	1520
1540	1600	1600	1710	1720	1720	1756	1760
1760	1800	1830	1860	1890	1910	1911	1950
1970	1982	2010	2050	2150	2240	2325	2328
2350	2380	2430	2490	2594	2620	2679	2945
2962	3200	3540	4330				

STATION		COUNTRY		RIVER		LOCATION	
CG60		FINLAND		KYMIJOKI		PERNOO	
138	159	183	233	258	263	270	290
308	312	312	320	338	342	342	343
347	357	357	366	367	385	385	388
391	393	406	412	415	416	418	435
436	445	454	458	463	467	471	471
472	474	494	507	507	508	512	517
520	527	527	535	537	540	542	546
547	552	557	558	563	574	578	584
584	614	616	644	658			

STATION	COUNTRY	RIVER	LOCATION
CG81	FINLAND	VUOKSI	IMATRA

333	341	408	448	461	461	476	479
491	497	506	508	534	534	540	548
561	582	590	599	603	603	604	605
607	613	616	624	630	636	639	642
642	642	651	651	651	656	658	659
659	664	666	668	673	677	677	677
680	684	686	686	686	686	689	691
702	703	703	703	703	703	706	710
712	712	718	721	721	727	727	727
730	730	736	739	742	744	744	744
744	747	756	759	760	760	766	766
769	773	775	788	789	792	792	793
794	794	795	799	803	806	818	829
836	839	840	846	864	880	882	887
911	914	917	928	936	1099	1109	1137
1146	1170						

STATION	COUNTRY	RIVER	LOCATION
BF42	POLAND	ODER	GOZDOWICE

707	726	733	799	828	830	850	860
866	885	906	915	920	947	970	975
978	1070	1080	1110	1140	1160	1160	1170
1200	1210	1210	1240	1240	1300	1300	1320
1320	1350	1360	1370	1400	1400	1430	1470
1550	1590	1620	1660	1690	1700	1710	1710
1740	1740	1800	1810	1830	1860	1930	2070
2140	2180	2290	2380	2420	2450	2480	2650
2980	3340	3720					

STATION	COUNTRY	RIVER	LOCATION
OF28	SWEDEN	VANERNGOTA	VANESBORG

353	355	399	405	407	419	419	421
450	454	455	462	467	473	475	477
481	481	487	487	492	492	494	494
500	500	504	504	508	510	512	515
516	518	527	529	535	535	537	539
539	541	551	551	551	552	553	557
564	564	568	568	570	574	578	580
582	582	584	584	585	588	588	590
592	592	592	593	597	599	601	601
603	607	609	610	615	619	621	623
625	628	630	632	634	634	636	636
637	640	642	642	644	644	645	646
648	648	648	652	654	658	660	663
669	671	671	671	672	673	673	675
677	677	677	681	691	693	693	694
702	706	706	708	712	716	718	722
726	726	728	731	731	735	737	739
743	743	745	751	759	761	768	768
772	772	774	774	774	780	782	784
794	798	817	817	829	836	926	928
938	1033						

STATION	COUNTRY	RIVER	LOCATION
DF09	USSR	NEVA	NOVOSARATOVKA

2000	2300	2500	2600	2650	2700	2700	2700
2700	2700	2700	2800	2800	2800	2800	2800
2800	2800	2900	2900	2900	2900	2930	3000
3000	3000	3000	3000	3000	3000	3000	3040
3100	3100	3100	3100	3100	3100	3100	3100
3100	3100	3100	3200	3200	3200	3200	3200
3200	3200	3200	3240	3300	3300	3300	3300
3300	3300	3300	3320	3400	3400	3400	3400
3400	3400	3400	3400	3400	3400	3400	3440
3500	3500	3500	3500	3500	3500	3500	3600
3600	3600	3600	3600	3600	3700	3700	3700
3800	3800	3800	3800	3800	3900	3900	3900
4000	4000	4000	4000	4000	4000	4000	4100
4100	4200	4300	4500	4500	4600		

STATION	COUNTRY	RIVER	LOCATION
JE9955	CANADA	ASSINIBOINE	BRANDON

22	23	37	39	47	49	70	71
74	75	77	81	85	88	90	90
94	99	99	103	105	112	116	116
120	133	134	134	135	145	146	151
154	157	159	159	165	165	166	174
184	187	199	202	210	212	214	217
222	229	241	243	258	258	303	314
360	360	422	430	450	484	541	603
651							

STATION	COUNTRY	RIVER	LOCATION
JE791	CANADA	RED	EMERSON

121	136	141	155	165	179	190	206
213	223	225	227	312	326	348	362
379	388	394	411	413	433	445	476
496	496	510	535	535	544	569	581
589	663	683	685	725	733	736	753
756	787	790	804	827	833	835	864
940	943	954	1120	1310	1310	1470	1550
1880	2670						